

AIDS myths that need eradicating

As emerged vividly at this year's international AIDS conference, prospects have never been so bright for new therapies. But some tales are spreading that are fictitious.

SCIENTISTS, activists and even journalists at the eleventh international conference on AIDS in Vancouver, Canada, earlier this month (see *Nature* 382, 101; 1996) spent much of their energy trying to ensure that the progress made so far in fighting the disease is kept in perspective. Initially, their effort paid off and most of the tidal wave of media coverage from the meeting was reasonable and fair. But as time passes, memories of that perspective fade, and some outlandish versions of events are gaining currency. There are several myths that need to be repudiated.

The largest states that combination drug therapies now under small-scale trial can be regarded as a likely 'cure' for HIV disease. The 'c' word was assiduously avoided by drug company scientists and other experts at Vancouver, but its misuse persists.

The trials undertaken so far of various promising combinations of protease inhibitors and other therapies are too small and too idealized to assure that the treatments will permanently reduce viral load. Even if they do, there is no evidence (or likelihood) that they will eradicate the virus from the body; the chances are that it will remain in the lymph nodes and in the central nervous system, requiring indefinite treatment with powerful drugs whose long term health effects are unlikely to be entirely benign. The success of the combination therapies in early trials is an encouraging sign on an otherwise bleak landscape; it is not a cure.

An opposite misconception of the combination therapies holds that they constitute little more than a plot by drug companies to extract money from the victims of the disease. The cost of the therapies — at \$6,000 or more a year for each component of the three-way combinations — is seen by many as not just unaffordable, but morally offensive. Larry Kramer, the respected AIDS activist, recently attacked the pricing in the *New York Times*, calling for "a programme to coordinate research, contain drug company greed and insure us all adequately".

An argument can certainly be made that capitalism is morally bankrupt. But it is profit-driven competition between corporations that has led to the protease inhibitors that now bring AIDS sufferers some hope. The divisions of those corporations that are engaged in pertinent research need revenue now, in order to win resources from inside the corporations for the further research which will duly yield better drugs. This mechanism may lack a moral dimension — but no known alternative model will yield better results.

The third myth surrounding the combinations is that they will chiefly serve to mutate the virus into new, and even more drug-resistant forms. But the single most promising aspect of the combinations is their apparent ability to restrict the number of mutations sufficiently to prevent this from occurring. The argument that diseases should not be fought because they then become more resistant is an argument for the cessation of all research into their cause and treatment.

A fourth myth contends that the character of the AIDS research community has undergone a fundamental change in the last two or three years, spawning new promise in the field. That fairy tale holds that a new, non-territorial, non-combative class of AIDS researcher has emerged, motivated only by the greater good, and free from the factionalism which supposedly consumed the 'old guard' after they began the war against the

disease a dozen years ago.

The US research effort may have shifted from the laboratories of the National Institutes of Health to new, extramural teams. But the 'new guard' has its factions and its egos too, and the old guard deserves some respect. Attempts by the new generation to contrast itself with the last one suggest a dangerous degree of self-delusion.

The final Vancouver myth holds that the arrival of effective drug therapies for use in the developed world will destroy the prospects for an effective vaccine for the developing world, where nine-tenths of new HIV infections now arise. That could happen if such drug treatments become available, but it need not. The community has been warned of the danger, and must act accordingly.

Ultimately, the virus will only be eradicated by a vaccine. Several initiatives are now under way to clear obstacles which lie in the path of vaccine development. Researchers and research funding institutions in every part of the world must ensure that this goal is kept firmly in mind. □

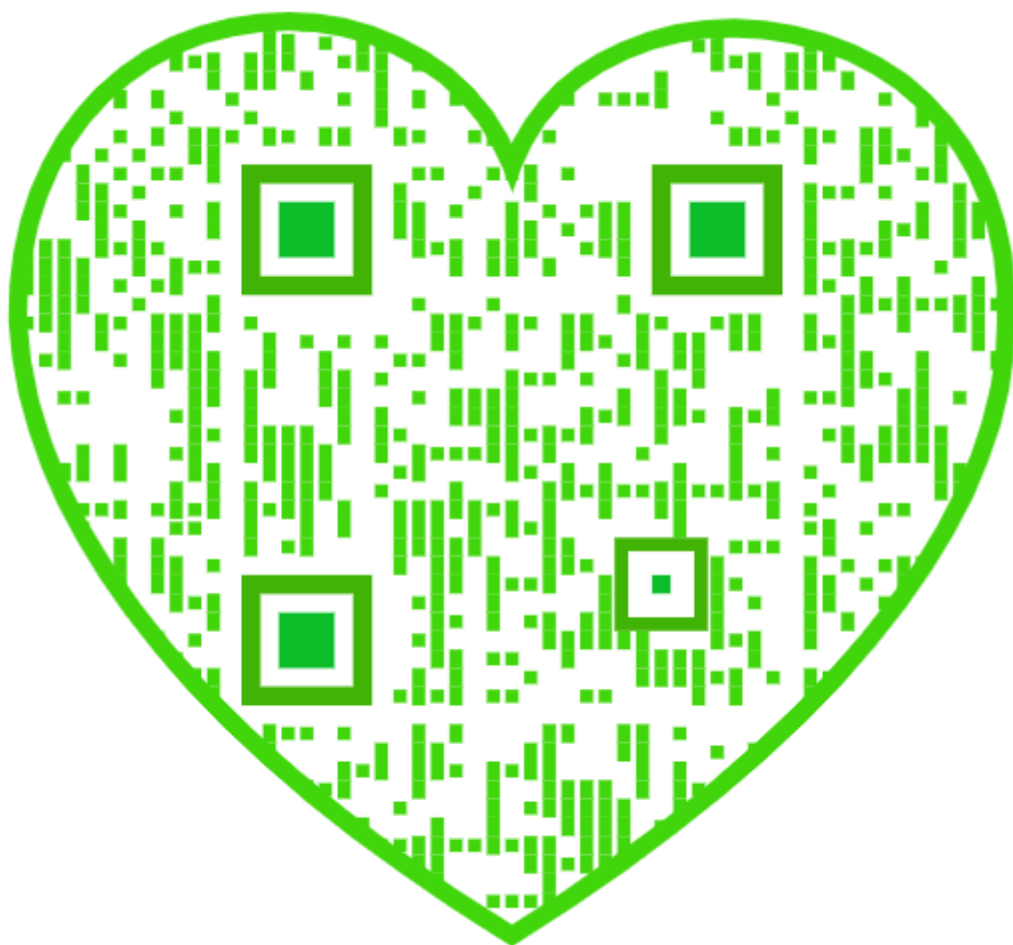
German lessons

The German government is to reduce its contribution to European laboratories. Scientists need to be realistic.

GERMANY'S tough squeeze on public spending announced earlier this month (see *Nature* 382, 196; 1996) implied cutbacks in support for applied research programmes, non-renewal of some research projects, and protection in general for basic science. But a third strand has now come to light: cuts averaging 6 per cent are to be imposed on German contributions to five major European research institutions (see page 285). These include the European Laboratory for Molecular Biology, the European Synchrotron Radiation Facility, and — perhaps most prominently given the current delicate state of negotiations over the Large Hadron Collider — the European Laboratory for Particle Physics (CERN).

The cuts will come as little surprise to close followers of European science policy. Evidence has been growing in recent years that the country's previous willingness to help pay for large scientific projects can no longer be taken for granted. But the new proposals will still come as a shock to the many scientists they will inevitably affect. Although the cuts have still to be formally approved by the Bundesrat, Germany's higher legislative body, that looks unlikely to represent a major hurdle.

Two conclusions can be drawn. The scientific community will be deluding itself if it fails to accept that the shift in Germany's thinking is part of an unstoppable trend that is part of the uncomfortable political and economic reality of the late 1990s. Second, all preconceptions need a critical reappraisal. For example, just how European does CERN have to be? Only after such questioning and some imaginative thinking will it be possible to negotiate structural adaptations that best protect the long-term interests of basic research. And only then will it be sensible and possible to correct the imbalance in commitments to national and international expenditures on basic science that many industrialized countries, not just in Europe, are now facing. □



German budget cuts cast chill wind over outlook for international labs

London. The German government is sending shock waves through Europe's scientific community by proposing to reduce its contribution to the support of five international research organizations by an average of six per cent in the next financial year.

The reduced subscriptions, the details of which are yet to be fully agreed, are a result of the harsh cuts in public spending announced two weeks ago by the conservative government of Chancellor Helmut Kohl as a step towards meeting the agreed criteria for membership of the single European currency (see *Nature* 382, 196; 1996).

While being presented by Germany primarily as a move designed to curb public spending and increase the efficiency with which research funds are spent, many scientists fear that it could help crystallize a general move throughout Europe to reduce spending on international basic science facilities, where these cannot demonstrate a likely contribution to the creation of economic wealth.

The European Laboratory for Particle Physics (CERN), in Geneva, is likely to be one of the hardest hit. Preliminary figures being circulated by the Federal Ministry of Education, Research and Technology in Bonn indicate that the government is proposing to reduce spending on CERN from DM265.7 million (US\$178 million) in 1996 to DM249.80 million next year.

Other bodies facing potential cuts in their German support include the European Molecular Biology Laboratory (EMBL) in Heidelberg, the European Southern Observatory, which is based in Garching and operates telescopes in Chile, and the Institut Laue Langevin (ILL) in Grenoble (see below). In percentage terms, the biggest reduction would be in Germany's contribution to the European Synchrotron Radiation Facility (ESRF), in Grenoble, which would fall by 10.2 per cent in constant terms.

Research ministers attending a private meeting recently of the industrialized nations belonging to the so-called G7 group are said to have discussed the growing problems many of them face in sustaining their subscriptions to international scientific organizations.

"It seems likely that, before long, a thorough analysis will be carried out of all major scientific projects to determine which are really necessary now, and which could be postponed to release funding for social programmes," Boris Saltykov, Russia's minister for science and technology policy, told a press conference in Moscow

shortly after returning from the meeting.

According to Saltykov, international projects are a particular target, as taxpayers do not understand why a substantial amount of their money is being spent abroad at a time when scientists are not receiving enough support at home. Germany, it appears, is no longer immune to this dilemma. Indeed, Jürgen Rüttgers, the country's research minister, is said to have described to his fellow G7 ministers a growing feeling in Bonn of the need to replace the country's previous largesse with a more hard-headed approach.

The implications of the proposed cuts are already sending a chill around the five international organizations concerned. This is particularly true since, in most cases, Germany's share of their budget is determined by a legally-binding treaty, and could not be reduced without accepting a comparable reduction in the contributions from other member states.

Officials in Bonn say that, in the short term, they expect to find ways to ensure that their commitments are met. In the long-term, they add, their goal is primarily to make German researchers more closely involved in decisions about distributing research funds between competing projects.

But Yves Petroff, for example, the director of ESRF, points out that if Germany's proposed move is copied by other member states, the result could be a major financial blow to the ESRF's research programmes, as much of its budget is already committed to salaries and basic running costs. "That could even mean that we would have to shut down for about six months," he says.

It would also make it impossible to complete the facility's remaining beam-lines. If that were to happen, he says, it would be

"extremely detrimental" to the ESRF. "Obviously from my point of view it is a very serious problem, and I am very worried."

There are similar concerns at the potential implications of the proposed reductions among German scientists. "The type of international collaboration represented by CERN is exemplary," says Volker Soergel, a former director of the DESY Research Centre in Hamburg. "I would hope that the people in Bonn would understand that if they are trying to develop new approaches."

But British officials, while expressing their own reservations about the possible impact of substantial cuts if not handled sensitively, are keen to point out that the German move indicates that they are no longer isolated in wanting to re-appraise how international scientific facilities are funded.

During a parliamentary debate last Friday, for example, Britain's science minister, Ian Taylor, said that now other countries "are becoming concerned about budgets", there was an opportunity to "work closely with the Germans" on the difficulties both countries faced in finding the money to support the construction of the LHC while at the same time maintaining a healthy domestic physics programme.

"It is absurd for us to have budgets that mean that we are effectively funding others to do research because we do not have enough left in our own budgets to carry out our national programmes, or get involved nationally in some of the work at the LHC," Taylor said.

Earlier, in a letter to Peter Williams, the chairman of the Particle Physics and Astronomy Research Council, he pointed out that, as a result of recent changes in voting rules, Britain and Germany can jointly block any future spending plans for CERN if they so choose. But science policy officials in other countries, faced with their own funding problems, indicate that Britain and Germany are not alone in seeking alternative, possibly drastic, solutions.

David Dickson, with Carl Levitin & Quirin Schiermeier

(see also pages 288 and 291)



Rüttgers: period of largesse is over.

Proposed reductions in German contributions to European laboratories (DM million)			
	1996 actual	1997 planned	reduction
European Laboratory for Particle Physics (CERN)	265.7	249.8	- 6.0%
European Molecular Biology Laboratory (EMBL)	22.2	21.2	- 4.5%
European Southern Observatory (ESO)	38.6	37.5	- 2.8%
European Synchrotron Radiation Facility (ESRF)	30.3	27.2	- 10.2%
Institut Laue-Langevin (ILL)	32.9	30.8	- 6.4%
TOTAL	389.7	366.5	- 6.0%
Source: Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie			

Rival Parkinson's bills focus US debate on fetal tissue

Washington. A struggle is developing in the US House of Representatives between two bills that would both substantially boost funding for research on Parkinson's disease, but whose differences highlight the continuing political battle over whether fetal tissue should be used in such research.

One bill, named after former congressman Morris K. Udall, would raise funding by the National Institutes of Health (NIH) to \$100 million in 1997, from the current level of \$28 million, and set up ten Parkinson's disease research centres. The other, introduced by Chris Smith (Republican, New Jersey), is virtually identical but bans federal funding for Parkinson's research using fetal tissue from induced abortions.

NIH officials, led by Harold Varmus, the director, have argued against the Udall bill on the grounds that they could not "responsibly" make use of such a sudden, large increase in Parkinson's research funding, and that generic nerve cell research will produce advances in fighting the disease.

At present, two studies sponsored by the National Institute of Neurological Disorders and Stroke (NINDS) involve transplants of neural tissue from voluntarily aborted fetuses to Parkinson's patients. The studies, at the University of Colorado in Denver and Mount Sinai Medical Center in New York, involve 76 patients and together received \$2.3 million in 1995.

In 1993, Congress passed and President Bill Clinton signed an NIH reauthorization bill lifting a Reagan and Bush-era ban on government funding of research using fetal tissue transplants from induced abortions. Advocates for Parkinson's research argue that such experiments represent a transitional phase, and that alternative sources of cells for transplant, such as pig fetal tissue, will be widely available within a few years.

The Udall bill has been introduced by Henry Waxman (Democrat, California) and has 209 co-sponsors, an unusually large backing in the 435-member House. These include liberal Democrats and prominent conservatives such as Thomas DeLay (Republican, Texas), House Majority Whip, and John Kasich (Republican, Ohio), chairman of the House Budget Committee.

The bill has 54 co-sponsors in the Senate, where it was introduced by Mark Hatfield (Republican, Oregon), chairman of the Appropriations Committee. Last week, the Senate Labor and Human Resources Committee unanimously incorporated it, in modified form, into legislation reauthorizing spending on the NIH to the end of the century. As passed by the committee, the NIH Revitalization Act authorizes \$80 million for Parkinson's research in 1997.

Paul Wellstone (Democrat, Minnesota), the chief committee backer of the provision, and both of whose parents had the disease, called the committee's action "an enormous step forward". But the bill's momentum is being blocked in the House, where abortion politics appear to be playing a large role during an election year. Michael Bilirakis (Republican, Florida), chair of the subcommittee on health of the House Commerce Committee, has not taken action on the Udall bill, which was referred to his subcommittee in April.

A spokesman said that Bilirakis, whose youngest brother died of the disease, is "very concerned" about the bill and "definitely supports [it]", but that a packed legislative calendar has prevented him from taking action. But others maintain that politics are behind the bill's stalled progress. Bilirakis is "essentially blocking it from getting to the floor," said Joan Samuelson, president of the Parkinson's Action Network, an advocacy group.

Waxman's spokesman Phil Schilero called the House inaction "very unfortunate". He said "too many people in the country want this to have Congress ignore it," reflecting the fact that about 1.5 million Americans suffer from Parkinson's. Earl Hilliard (Democrat, Alabama), is gathering signatures for a letter to Commerce Committee chairman Thomas Bliley (Republican, Virginia), complaining about Bilirakis' inaction.

Parkinson's researchers are also disappointed. Abortion "is a public policy issue and shouldn't be decided in a scientific arena where scientists are seriously attempting to help patients with horrible disabilities," said Warren C. Onslow, the principal investigator on the Mount Sinai grant.

Smith's alternative bill has gained 15 co-sponsors. Smith circulated a letter to House members on 12 July outlining current NIH projects involving fetal tissue transplants and urging them to switch support to his bill.

Cliff Stearns (Republican, Florida) is among several conservatives who have removed their names from the Udall bill. He told the House last week that he did so "due to my concerns that it allows the NIH to expand its research using tissues from aborted babies".

Zach Hall, the director of the NINDS, which funds most NIH research into Parkinson's, supported the use of fetal tissue as a transitory step in testimony last month to a subcommittee of the Senate Appropriations Committee. No-one believes that fetal tissue transplants will be carried out on a large scale in ten years, he said. "We view this as a step in the research and not as an end."

Meredith Wadman

New chairman sees 'outward-looking' role for science board

Washington. The new chairman of the US National Science Board (NSB) has promised that it will "look more outwards" from its immediate task of governing the National Science Foundation (NSF) to offer advice on key science policy questions.

Richard Zare, 56, a professor of chemistry at Stanford University, chaired his first board meeting last week. He says that during his two-year tenure he would like the board to deal with issues such as the "proper role of the federal government in research", the nature of the connection between research and development, and ways of handling international scientific collaborations.

The NSF funds most non-biomedical university research in the United States. Zare



Zare: risk-aversion may have its dangers.

says that the board's priority in governing the NSF will be to ensure that merit review works fairly and that the foundation allows enough emphasis on emerging disciplines.

"The question is how do you know when to stop good work in order to start better work that might fail?" he says. Zare sees a danger that agencies such as NSF may become reluctant to take risks now that money is tight. He concedes that taking the "high ground" and pursuing new work will be "politically difficult and may cost the NSF support in the Congress".

Zare says that the board will look at revising the format of *Science and Engineering Indicators*, the NSF's information-packed but sometimes impenetrable biennial report on the state of US research.

NSB members are appointed by US presidents to serve rolling six-year terms. The whole panel is therefore not picked by the sitting president, who usually turns instead to his own panel for advice — the President's Council of Advisors in Science and Technology, in Bill Clinton's case.

As Zare arrived in his part-time position, the NSF announced the departure of its deputy director Anne Petersen after less than two years, to join the Kellogg Foundation, a Michigan-based charity. The loss of Petersen, responsible for much of the day-to-day running of NSF, comes at a bad time for the agency, which may face severe staff cuts this autumn (see *Nature* 382, 7; 1996).

November's presidential election makes it unlikely that a permanent replacement will be confirmed before the beginning of next year.

Colin Macilwain

United States backs climate panel findings

London. The US government last week condemned attempts by US energy interests to undermine the work of the Intergovernmental Panel on Climate Change (IPCC). Committing itself for the first time to binding targets to reduce greenhouse gas emissions, it said the science of climate change "cannot be ignored" and that human-induced global warming "cannot be wished away".

Timothy Wirth, US Under-Secretary of State for Global Affairs, said the findings of the IPCC had convinced the Clinton administration of the need to take steps to address global warming. Wirth was addressing the second annual conference of the United Nations climate convention in Geneva.

Wirth described as a "remarkable statement" the conclusion of the IPCC's latest report on climate change, that "the balance of evidence suggests that there is a discernible human influence on global climate". He said the administration took the report "very seriously".

Wirth described the IPCC's critics as "naysayers and special interests bent on belittling, attacking and obfuscating climate change science". John Shlaes, executive director of the Global Climate Coalition, the lobby group that has tried to discredit parts of the IPCC report, said the statement would cost the US "millions of jobs".

Wirth's speech initially delighted environmentalists at the Geneva conference. It also set the tone for an upbeat statement in which most of the 60 environment ministers present promised to start negotiations in December on an international law to bind countries to reducing carbon dioxide (CO₂) emissions by a specified quantity after 2000.

But a closer reading of the speech also reveals the Clinton administration's terms for future US cooperation in climate negotiations. Wirth said that all present proposals to limit CO₂ emissions were "neither realistic nor achievable". The Alliance of Small Island States (AoSIS) would like a 20 per cent reduction in CO₂ emissions, compared

to 1990 levels, by 2005, and Britain has suggested a 5–10 per cent cut by 2010.

But environmentalist organizations, such as the Climate Action Network (CAN), say they will continue to push for the AoSIS proposal as the basis for future reductions targets. "Tactically, it's our best option," says Gurmit Singh, regional coordinator for CAN south-east Asia. "If we start with a weaker text, it's bound to get weaker still by the time it's agreed."

Wirth also emphasized that the United States would oppose any law binding countries to specific measures to reduce CO₂ emissions. Furthermore, a long-term international agreement to reduce CO₂ emissions, he added, would need to include commitments from developing countries.

The speech angered Russia, Australia and ten states belonging to the Organization of Petroleum Exporting Countries (OPEC). An OPEC statement said its members would lose US\$190 billion in oil export income if a carbon tax were introduced.

Developing countries also left Geneva concerned at the implications that CO₂ reduction measures would have on their newly industrializing economies. The current climate convention asks that only developed countries reduce their CO₂ emissions to 1990 levels by 2000.

Li Zhaoxing, China's Vice-Foreign Minister and head of the country's delegation to the climate convention, said his government attaches "great importance" to climate change. But he also added that "China is a developing country whose first and overriding priorities are economic and social development; to eradicate poverty and to meet the basic needs of the people's livelihood".

Michael Grubb, head of the energy and environment programme at the Royal Institute for International Affairs in London, says that designing an agreement that includes developed and developing

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Culture clash? China's development priorities seem at odds with calls to limit carbon dioxide emissions.

countries "will be an extremely tough issue". The US government, says Grubb, understands it is inequitable to ask developing countries to reduce their CO₂ emissions at a level comparable to the developed world. But he adds that the Clinton administration could not risk the potential political backlash at home if developing countries remained exempt.

Jake Werksman, of the Foundation for International Environmental Law and Development at the University of London, believes that the US government will "most likely" insist that developing countries sign up to an agreement that allows them to defer emission-reducing measures. "No developing country will be willing to sign up to anything stronger," he says.

The Wirth speech, an apparent reversal of US climate policy, is believed to reflect the growing confidence within the Clinton administration of victory at the November elections. But Rafe Pomerance, deputy assistant secretary for environment and development at the State Department who helped draft the statement, says the timing of its release is a coincidence. "The administration has been working on the policy for more than a year."

The speech caused further controversy over the absence of the word 'legal' before references to 'binding targets' to limit CO₂ emissions. Singh and Grubb both feel that its absence is significant. But, whereas Grubb believes the omission is no more than "political sleight of hand" designed to pacify a potentially sceptical American public, Singh regards it as an "escape clause" designed to protect the United States from signing up to any law that could cause political difficulties at home.

Mark Hambley, a senior official with the US delegation to the Geneva talks, says the word's absence is "not an issue". He points out that the US government "wholeheartedly endorsed" the end-of-conference ministerial declaration that calls for "quantified legally-binding objectives for emissions limitations".

Ehsan Masood

Lobbyists 'belittle climate change science'

London. Timothy Wirth, US Under-Secretary of State for Global Affairs, opened his address to last week's climate conference (see above) with a strong defence of the Intergovernmental Panel on Climate Change (IPCC) against the "false" charges of scientific impropriety.

The US government was "not swayed by and strongly objected to" allegations made by groups such as the Global Climate Coalition (GCC) about the integrity of the IPCC's conclusions (see *Nature* **381**, 546; 1996), Wirth said. "Let's take a false issue off the table. There can be no question but that the

IPCC's findings meet the highest standards of scientific integrity".

Ben Santer, an IPCC lead author and atmospheric physicist at the Lawrence Livermore Laboratory, California, who is at the centre of the controversy over the rewriting of a key chapter, says the statement amounts to "a ringing endorsement" of the IPCC.

Dana Rohrabacher, chair of the House Science Committee's energy and environment subcommittee, is considering convening congressional hearings at which Santer would be asked to explain his actions.

E. M.

Funding forces fusion project out of Europe

Munich. The International Thermonuclear Experimental Reactor (ITER), a multi-billion dollar project to assess the feasibility of controlled nuclear fusion, will almost certainly not be built in Europe. This follows a joint announcement last week by the French and German research ministers that they are withdrawing their interest in offering sites for ITER, as the projected costs for the host country had risen too high.

The ministers expressed continued support for ITER and for fusion research in general. But their announcement throws into further doubt the future of a project that has already suffered a series of blows over the past year, particularly a decision by the United States to pull back sharply from its commitment to fusion research (see *Nature* 379, 387; 1996).

Nevertheless, the Franco-German announcement came as a surprise to both fusion scientists and ITER administrators, as bids for the siting of the reactor are not yet due. ITER is in the middle of a six-year engineering design phase, scheduled for completion by 1998. A decision on where it should be built is not required until the end of next year, while negotiations on cost-sharing will begin in the next 12 months.

Given the current stage of discussions, Klaus Pinkau, director of the Max Planck Institute for Plasma Physics in Garching, says he was "astonished". It was a mistake for Germany and France "to show their cards before negotiations have started".

The German and French ministries of research declined to comment on the timing of their joint announcement. But it is widely

believed to have been taken as a result of pressures in both countries to reduce future expenditure in order to meet the 'Maas-tricht' criteria for joining the planned single European currency.

ITER is a joint venture of the European Union (EU), Japan, the United States and Russia. Problems with financing have arisen partly because Russia and the United States have said that they are no longer in a position to make the level of contribution envisaged when the project was conceived in the mid-1980s. As a result, the EU and Japan would have to pay a much greater percentage of the costs to host ITER, whose total bill will probably exceed US\$15 billion.

In an informal proposal to the ITER management earlier this year, Japan raised the stakes by saying that it would be prepared to pay all the basic costs of the project, if the reactor were to be built in Japan (see *Nature* 380, 655; 1996). It envisaged that the remaining costs, the so-called 'high-tech core', which it estimated at 40 per cent of the total, would then be split equally between the four partners. This would mean Japan paying 70 per cent of the total.

At the beginning of May, the European Commission (EC) distributed to EU member states a paper intended to help them prepare proposals for hosting the reactor. Germany was expected to propose a site at Greifswald, where its next-generation fusion machine, Wendelstein 7X, is being built. France was expected to propose a site at Cadarache, where the Commissariat à l'Energie Atomique (CEA) has its fusion facility.

Europe would like the shared high-tech

core costs to be much higher than that suggested by Japan, namely 70 per cent of the total. So the EC paper suggests that the total cost to the host country should lie between its own estimate of 47.5 per cent and the Japanese estimate of 70 per cent. The Franco-German statement, which was a response to this paper, does not indicate whether the two countries would be prepared to reconsider their withdrawal should the cost to the host country end up, after negotiations in the next two years, at the lower end of this scale.

Ernesto Canobbio, a commission official with responsibility for the ITER project, says that the timing of the statement is particularly surprising, as it also pre-empted the report of an independent committee set up by the commission to assess the international fusion programme. That committee's report is due later this year. Sergio Barabaschi, an industrial scientist and adviser on science to the Italian government who is the chairman of the committee, says that the timing of the Franco-German statement "does not follow any logic".

But Canobbio points out that early knowledge of France and Germany's decision makes life simpler. "At least we don't have to keep alive a fiction," he says. He now plans to negotiate the best possible deal for European researchers — assuming that the reactor will indeed be built elsewhere. Japan is now the only realistic option.

Pinkau says that the most important thing now is for ITER to go ahead "because it is the main thrust of all fusion research". Michel Chatelier, deputy head of the CEA Department of Controlled Fusion in Cadarache, says that scientists have always been aware that the high cost of ITER makes it vulnerable. He is relieved that the Franco-German statement signalled strong support for fusion research in general.

Alison Abbott

Russian contribution for CERN collider

Moscow. **Russia's participation in the planned \$3-billion Large Hadron Collider project (LHC) at the European Laboratory for Particle Physics (CERN) in Geneva has been confirmed in a decree published by the government last month. The government has agreed to contribute US\$60 million towards the design and construction of the LHC. In addition, 430 Russian scientists will work on the ATLAS, CMS and ALICE experiments.**

Russia has already completed the first of 65 modules for a hadron calorimeter for the ATLAS experiment on colliding proton beams. Scientists at the Joint Institute for Nuclear

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Research in Dubna, near Moscow, assembled and delivered the 6-metre, 20-tonne module (above) to CERN in April. Russia has been an active participant in CERN for the past four decades, and currently sends 1,000 scientists to the laboratory — the same number as the United States. □

Budget for US physics

Washington. **Appropriations committees of both the US Senate and House of Representatives have passed 1997 spending proposals for the Department of Energy that would maintain most physics research programmes just above this year's levels.**

There are slight reductions for magnetic fusion research, to which the House proposes allocating \$209 million and the Senate \$225 million in 1997, compared to this year's spending of \$227 million and the president's request of \$264 million. The House is also calling for drastic reductions in administrative spending at the department, including cuts of two-thirds in the public and congressional relations department, and seven out of eight members of the policy staff. □

Bollworms chew hole in gene-engineered cotton

Washington. A large question mark is hanging over the effectiveness and safety of one of the first genetically-engineered crops to be extensively planted, after cotton bollworms were found to have infested thousands of acres planted with the new breed of cotton in Texas.

Two million acres of so-called 'Bt cotton' have been planted for the first time this year, after the crop was approved last autumn by the US Environmental Protection Agency (EPA). The crop is named after the soil bacterium *Bacillus thuringiensis*, genetic characteristics of which are introduced into the cotton to produce a natural insecticide.

Scientists from the chemicals company Monsanto, which has invested millions of dollars to develop the crop, are this week collecting samples of planted seeds and bollworms to investigate the apparent breakdown in the crop's effectiveness.

Shares in Delta and Pine Land, which distributes the crop for Monsanto, were briefly suspended on the New York Stock Exchange on 15 July as news of the problem spread to Wall Street. Trading resumed at \$26, down from \$33 before the suspension.

The Union of Concerned Scientists (UCS), which has opposed certain applications of genetic engineering, called on the EPA to suspend registration of the sale of Bt cotton. But the EPA is waiting for data from Monsanto's scientists before deciding how to respond. Lynn Goldman, assistant administrator of the EPA, says: "We hope the issue can be resolved within weeks. But if Bt cotton doesn't work, it doesn't matter if we give them registration or not — they [Monsanto] aren't going to have a product."

With two million acres planted from Georgia to Texas — out of a total of about 15 million acres under cotton in the United States — Bt cotton is perhaps the most economically significant of all genetically engineered plants introduced in the country.

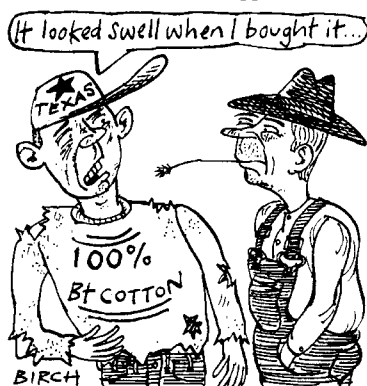
The genetically engineered cotton generates a natural toxin to kill caterpillars of three pests: cotton bollworm, tobacco budworm and pink bollworm. Critics fear that it may produce a generation of insects with resistance to the natural toxin. These would then attack the crops of organic farmers who rely on spraying with the natural toxins.

To avoid this risk, the EPA only allowed Bt cotton to be planted subject to a resistance management plan. This requires that the dose of toxin in the plants is high enough to kill almost all insects exposed to it, so that few survive to breed. Farmers must also keep some non-engineered crops nearby, to retain a population of non-resistant insects. Insects that do survive natural toxin will then interbreed with 'normal' insects, preventing

the development of a super-resistant strain.

Margaret Mellon of UCS says: "The cotton's failure to control bollworm indicates that the current management plan will not work for at least one pest. EPA should suspend the registration and bar sales of Bt cotton until the company develops a new resistance management plan that will work."

Monsanto says, however, that the plan works differently for different pests. Randy Deaton, the company's manager in charge of Bt cotton development, says that the cotton bollworm is unlikely to develop resistance, because it lives on so many plants other than cotton. "The biggest concern for



resistance is the tobacco budworm," he says. The budworm is killed by a lower dose of the natural toxin, and there is no evidence that it is surviving in Bt cotton.

According to Randy Luttrell of the agriculture school at Mississippi State University, the Mississippi farmers who paid a \$32-per-acre licence fee for the crop did so primarily for protection against tobacco budworm. But some farmers in the state are spraying Bt cotton as a precaution, he says.

Deaton says: "Growers have a natural anxiety this time of year. They'd be on edge no matter what." He declined to comment on rumours that insecticide suppliers are spreading stories about Bt cotton's failures to discredit the genetically-engineered crop, which would need no spraying if it worked properly.

Goldman says that the EPA will get data this week from Monsanto on whether the surviving bollworms are of a resistant strain. The genetically-engineered plant includes a marker gene to enable scientists to establish if infested plants are Bt cotton, or if a mix-up with seeds has contributed to the problem.

Monsanto hopes that the efficacy of its cotton plants has merely been limited by the severe drought in Texas, and that farmers will soon regain confidence in Bt cotton. "It has lots of benefits for growers and lots of benefits for the environment," says Deaton.

Colin Macilwain

Rights issue take-up for British Biotech 'disappointing'

London. The largest single fundraising attempt by a UK biotechnology company met with mixed reactions last week when only 49 per cent of shareholders in British Biotech took up their rights to buy additional shares.

Small biotechnology companies in the United States and United Kingdom raised almost £2.11 billion (\$3.3 billion) in public offerings in the year to 11 July, so the low take-up for the rights issue was a disappointment to British Biotech. But not a disaster, because underwriter Kleinwort Benson is to pay any shortfall in the £143 million anticipated.

Launched in 1986 and now valued at £1.4 billion, British Biotech has yet to bring a product to market, but is pinning its hopes on lexipafant, a treatment for pancreatitis, and the anticancer drug marimastat.

When promising results from phase I and II studies on marimastat were presented to the American Society of Clinical Oncology in May, the company's stock rose to £32.65. The drug appears to be active in patients with advanced pancreatic carcinoma, can be given orally and seems to be tolerated well. But the two-year phase III trials have only just begun. A spokeswoman for British Biotech is confident that "everything remains on track at the company" and there will be no need to generate further capital to bring marimastat to the market.

While the larger biotechnology companies remain fairly secure, new ones springing up in the wake of their success may face a struggle for funds.

Mike Ward, editor of *BioBusiness*, believes that the experience of British Biotech was not due to any "negative news" about the company, but rather that it comes at a time when too many companies are chasing a finite number of investors. "The well has run dry," he says. He admits that the risk inherent in the biotechnology sector may be best taken, or at least judged, by the larger pharmaceutical companies. Cantab Pharmaceuticals, one of the smaller companies, celebrated a dramatic rise in its share price last week on the announcement of its collaboration with Smith-Kline Beecham to develop a vaccine for genital warts.

There are concerns that relying on large companies to foster drug development means that treatments destined only for a limited market will remain unfunded. In the United States, companies may apply for 'orphan drug' status for such treatments, ensuring a monopoly position for the first seven years of marketing. The European Union is considering similar legislation.

Ginny Page

Brazilian researchers took first bite at 'new muscle'

São Paulo. A group of Brazilian researchers has been taken aback by the announcement by US dentists of the discovery of a new facial muscle, because they described it themselves in a local journal in 1978.

Gary Hack, assistant professor of dentistry at the University of Maryland in Baltimore, and Gwendolyn Dunn, an orthodontist, say they discovered the muscle, involved with mastication, because they opted for an unusual way of dissecting the face, carving not from the traditional direction, from the side of the head, but from the front (see *Nature Medicine* 2, 506; 1996).

The 'undiscovered' muscle was found to run from behind the eye socket to the lower jaw. Hack and Dunn went on to dissect other heads and to scan the scientific literature to find whether the muscle had already been described. They waited a year before making their announcement. "We reviewed 15 textbooks," says Hack, without finding the muscle.

The Americans called the muscle spheno-mandibularis, as it begins at the sphenoid and has an insertion point at the mandible. If correct, this would be the fifth muscle involved with mastication. But when the announcement was reported by the media, it came to the attention of Luis Roberto de Toledo Ramalho, a Brazilian researcher at the Faculty of Dentistry at UNESP (São Paulo State University) at Araraquara.

Toledo immediately remembered an

earlier research paper written with two colleagues, Carlos Landucci and Helio Ferraz Porciuncula, which was published in the *Revista da Faculdade de Odontologia*, the faculty's own journal. The journal is published in Portuguese and does not have a wide circulation.

Considering it unlikely that centuries of dissection would have let pass such a thing, the Brazilians had not thought that they were dealing with a new muscle, and they labelled it merely "deep portion of the temporal human muscle". Toledo sent a copy of his work, together with colour photographs, to Hack. "I am not interested in any sort of polemics," says Toledo, "I just want to show that we had already done work on this, and we do not believe it is a new muscle."

"We were fascinated," says Hack. "Their description goes beyond anything we have seen in any textbook. I wish we had known this work before. But we disagree about the conclusion. We are finding some very interesting data, including biochemical and physiological analysis, that shows it is a different muscle."

Whether or not Hack proves that the "deep portion" deserves to be called spheno-mandibularis, Brazilian scientists are sensitive to the fact that relevant research was buried in an obscure Portuguese-language journal. "I wonder what else could be hidden in these kind of journals," says Toledo.

Ricardo Bonalume Neto

Outbreak of *E. coli* infection in Japan renews concerns

Tokyo. Concern was sparked among public health officials around the world last week when more than 6,000 schoolchildren in Sakai, Osaka city, were struck down with infection by the 0157 strain of the bacterium *Escherichia coli*.

The strain was first linked with illness in humans in the United States in 1982, when it caused severe diarrhoea in individuals who had eaten contaminated hamburgers at a national fast-food chain.

The outbreak in Sakai follows smaller outbreaks in schools around Japan. These began in May in Okayama Prefecture, west of Osaka, where two children died of complications caused by a toxin released by the bacteria.

More than 600 of the children in Sakai have been in hospital and two are in a critical condition with cerebral bleeding and kidney failure. Another 75 have developed haemolytic uraemic syndrome (HUS), which is characterized by haemolytic anaemia and kidney failure.

Non-toxic strains of *E. coli* occur naturally in animals, including humans, helping to suppress the growth of harmful bacteria and synthesizing vitamins. The 0157 strain, which is carried by some cattle, is a rare variety that produces potent toxins.

Unlike other forms of food poisoning, where millions of bacteria have to infect a person to cause symptoms, it is suspected that, as in the case of dysentery, ingestion of as few as ten bacteria can cause disease. So secondary infections passed on via the stools of infected individuals are common. Children and the elderly are particularly susceptible to infection.

School lunches are thought to be involved in the current outbreak in Japan, although the source has been pinned down precisely in only one case; a young boy in Yokohama, near Tokyo, ate raw beef liver subsequently found to be contaminated with the bacteria. But preliminary DNA studies of bacteria from the various outbreaks, carried out by the National Children's Hospital and the National Institute of Health in Tokyo, suggest multiple sources for the outbreaks.

The Ministry of Health and Welfare is coming under severe criticism from health experts for not taking more stringent measures to contain the disease. The ministry has so far treated it as food poisoning rather than as an infectious disease. But many health experts argue that, because of the contagious nature of the bacteria and the common occurrence of secondary infections, the outbreaks should be covered by the Infectious Diseases Prevention Law. That would require quarantine and disinfection measures.

David Swinbanks

Cash and contacts are keys in the East

Munich. Basic science in the former communist countries of central and eastern Europe is hampered not just by an acute lack of money, but also by the difficulties of maintaining contacts between scientists, according to participants in a workshop in Paris last week, which drew together representatives from research organizations in 14 countries.

The workshop was sponsored by the International Council of Scientific Unions (ICSU), the European Union, the European Science Foundation, and INTAS, the European funding agency for scientists in eastern Europe.

The percentage of gross domestic product allocated to research has fallen dramatically in all the former Communist countries in the last few years, in some cases to only one-tenth of its 1989 value. In many countries new laws stipulate an increase in funding. But, said delegates, these are systematically

ignored by governments facing more urgent needs. What little money there is tends to be directed to applied research.

Communication between the new research agencies in such countries has also suffered since the fall of communism. The plethora of new agencies, with diverse responsibilities, is confusing to western research organizations, making international collaboration difficult. The workshop recommended the setting up of a home page on the World Wide Web to list and describe all such agencies in the West and the East.

The workshop asked for international support to help sift through the wealth of unpublished material in the archives of research institutes — particularly in the former Soviet Union — that were closed to the outside world during the communist era. It also asked for international support to promote the sharing of large research facilities among eastern European countries.

Alison Abbott

MPs back increase to fund hadron collider

London. A parliamentary committee has supported Britain's particle physicists in their efforts to persuade the government to provide an extra £60 million (US\$92 million) towards the construction costs of the Large Hadron Collider (LHC) at the European Laboratory for Particle Physics (CERN) in Geneva. The funds would be provided over eight to ten years.

In a report on the Particle Physics and Astronomy Research Council published last week, the House of Commons Select Committee on Science and Technology says that there is a "good case" for a general increase in the science budget. And CERN should receive "adequate funding" to exploit the scientific opportunities offered by the LHC.

But the report's conclusions were not unanimous. Several committee members voted for an alternative conclusion which stated that it would be "quite unrealistic" to expect the Treasury to provide the money being requested. They said that the solution "has to be sought in a reappraisal and re-deployment of the international efforts in research in particle physics".

The main argument of the report is that, in signing up with other CERN member states to the construction of the LHC at the end of 1994, Britain had agreed to a step that it can no longer properly afford at

current funding levels. The result is that the CERN subscription is becoming a subsidy of international science, "of very restricted value to British science".

A 25 per cent increase in the real costs of the subscription, at a time when the overall science budget is being kept level, will mean significant cuts in domestic expenditure on particle physics, including the development of LHC detectors.

The problems are caused largely by fluctuating exchange rates as the subscription is paid in Swiss francs. The committee says that, to avoid funding problems, such extra costs should be paid directly by the Treasury.

It also backs CERN's proposal that, if the government is not prepared to provide extra money towards the LHC costs directly, it should do so in the form of a loan (see *Nature* 381, 102; 1996). The research council's allocation would be reduced by an equivalent amount once the construction phase has ended, when CERN's annual budget is expected to

fall by about 20 per cent.

The report has been welcomed by particle physicists, who have been expressing concern at the potential consequences of meeting the LHC costs out of CERN's commitment to their field — exacerbated by the research council's decision to put a cap on this commitment at its current level (see *Nature* 380, 576; 1996).

But enthusiasm for the select committee's conclusions has been tempered by the fact that three of the nine members present when the report was approved voted for a different set of recommendations. The main thrust of these is that the solution can only be found by discussions between the science ministers of the world's top seven industrialised nations.

The alternative amendment was drafted by Jeremy Bray (Labour, Motherwell South), a former shadow research minister. This warned that, without increased participation in CERN — "or the reconstruction of CERN on a global rather than a European basis" — the LHC timetable "may have to be rephased more drastically than has yet been considered".

"CERN is faced with the need for additional funds, and the only practical way of doing this is for CERN to get increased income from other sources," says Bray.

David Dickson

Universal Pictorial

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Bray: timetable may need drastic revision.

Asbestos removal spells upheaval for Paris scientists

Paris. A surprise announcement by Jacques Chirac, the French president, that "there will be no students left at [the university campus of] Jussieu by the end of 1997" because of asbestos contamination, was refuted a few days later by François Bayrou, the education minister.

Chirac's announcement in a national broadcast on Bastille Day (14 July) caught everybody by surprise, including Bayrou, who promptly appointed a 'Jussieu Mission' committee. But four days later, Bayrou announced that there would be no relocation of the 40,000 students and 10,000 staff from Europe's largest science campus.

Nevertheless, he promised that "there would be no limit to the state's financial commitment" to remove the asbestos, which is contaminating 220,000 square metres of teaching and research facilities at Jussieu.

The Jussieu Mission includes representatives of the campus's two universities, Paris VI and Paris VII, as well as the Institute of Physics of the Globe. It will present a detailed plan in the autumn for decontaminating the campus, which was built between 1965 and 1972.

Faculty members had been alarmed at the prospect of relocating under such pressure. The presidents of the two

universities, Jean Lemerle (Paris VI) and Jean-Pierre Dedonder (Paris VII), declared that "closing down Jussieu will spell the death of Paris VI and Paris VII and deal a severe blow to French scientific research".

Bayrou described Chirac's statement as a "boost" in dealing with a public health issue. But Ramez Sami, a mathematician who belongs to the Jussieu Anti-Asbestos Committee, a pressure group that has been active for 20 years, was less positive. "I don't understand why after so much resistance, such unrealistic and revolutionary measures are suddenly being proposed," he said.

The minister said priority will be given to emergency work to stop the release of asbestos dust. This will involve covering ceilings, walls and technical access areas with airtight plastic sheeting "before the end of 1996, or at the latest during the first quarter of 1997". No specific budget has been allocated for this work, which is minor compared to removing the asbestos, estimated to cost of FF880 million to 1 billion (US\$176 million to US\$200 million).

A source at the Ministry of Education said that "the real cost will probably be closer to FF2 billion". She also doubted that the emergency phase of

the project would be completed by 1997.

It now appears that decontamination of the buildings will be done in stages, which will require only temporary shutdown of individual facilities while work is under way.

A senior member of the administration of the Institute of Physics of the Globe, Liliane Flabée, said that "we have already sealed off our ceilings and technical areas. This means we can lay no new cables or drill holes" because of the danger of releasing asbestos dust. "When the asbestos is removed," she predicted, "we will have to stop much of our research and we don't know for how long".

Research laboratories will be harder to decontaminate than teaching facilities. "Moving out massive equipment such as mass spectrometers, mainframe computers and computer networks is not going to be as easy as moving a desk and a PC," Flabée pointed out.

Michel Parigot, president of the Anti-Asbestos Committee at Jussieu, says that Chirac's pledge to tackle the Jussieu issue had deterred him from taking legal action against the government for negligence. He estimates the Jussieu clean-up to be the largest operation of its kind in France.

Ronnie Amelan

Approval closer for RU-486

Washington. An advisory committee to the US Food and Drug Administration (FDA) has concluded that the abortion-inducing drug RU-486, known in the United States as mifepristone, is safe and effective, and recommended that it be approved for marketing.

Although the recommendation of the eight-member panel — known as the Reproductive Health Drugs Advisory Committee — is not binding on the FDA, such decisions are rarely ignored. David Kessler, the FDA's Commissioner, said the panel had issued "a very strong signal". A final decision is expected in September. □

Staff 'negligent' over CJD risk

London. Officials and staff from Britain's Department of Health and the Medical Research Council (MRC) were negligent in not responding adequately to warnings about the risk of contracting Creutzfeldt-Jakob disease from human growth hormone in experimental treatments during the 1970s, the High Court ruled last Friday.

Having been alerted to the possibility of transmission, the department and the MRC failed to pass on the information to clinicians treating patients for dwarfism, and should have ended the programme before 1 July 1977. The MRC expressed its "deepest sympathy" for the families of the eight patients who died, and the many who may still develop the disease, but did comment on the verdict. □

German institutes face closure

Munich. Three of Germany's 83 'blue-list' institutes — the Research Institute for Child Nutrition in Dortmund, the Medical Institute of Environmental Hygiene in Düsseldorf, and the Institute for Scientific Film in Göttingen — face closure, following a recommendation from the Wissenschaftsrat, Germany's science council, that their funding be withdrawn.

Earlier this month, the Wissenschaftsrat, which is responsible for reviewing the research activities of all blue-list institutes, published reports on nine of them. It said that the three did not comply with the appropriate "scientific and political criteria", while the other six institutes covered received positive reviews. □

Clinton names bioethics head

Washington. President Clinton last week named Harold T. Shapiro, the president of Princeton University, to chair a new 15-member National Bioethics Advisory Commission, which will be concerned with ethical issues arising from experiments involving human biology and behaviour. First among the topics it is mandated to consider are protection for human research subjects and the use and management of genetic information, including gene patenting issues.

The commission includes experts in medicine, religion, psychology, public health and law, as well as lay people. Shapiro is a professor of economics and public affairs. The panel is intended to address broad principles, to guide federal agencies, to hold hearings, to lead inquiries and to issue recommendations. □

New director for Newton Institute

London. Keith Moffatt, (right) professor of mathematical physics at the Department of Applied Mathematics and Theoretical Physics at the University of Cambridge, has been appointed the new director of the university's Isaac Newton Institute for Mathematical Sciences.

Moffatt takes over from Sir Michael Atiyah, recently-retired president of the Royal Society. A specialist in fluid dynamics, he has been closely involved with the institute since its formation in 1990 and as a member of its management committee. □



Dona Haycraft

Does your automated DNA sequencer leave you guessing? If so, chances are it's primarily designed for high throughput sequencing. Why be uncertain of your sequencer's accuracy, when ALFexpress™ is providing researchers with the full genetic stories of their DNA.

ALFexpress: for more accurate readings

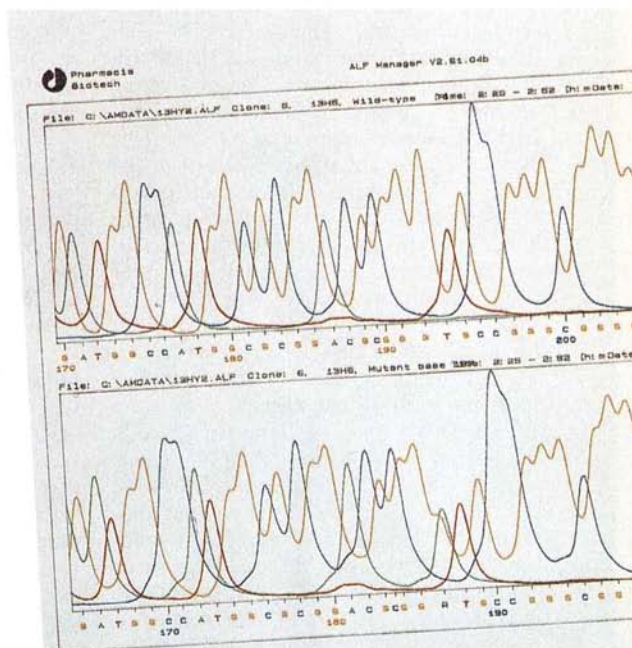
In the largest clinical study using automated DNA sequencing, the technology behind ALFexpress proved exceptional (see caption). That's one of many examples of ALFexpress offering unrivaled accuracy during automated confirmatory sequencing. Further, its readings are so accurate that ALFexpress can unambiguously identify heterozygous point mutations—as proven in many clinical research applications, such as analysis of tumor genes and high-resolution HLA typing. What's more, Pharmacia Biotech has dedicated software programs to support these applications.

For the full story, call us: 1 (800) 526-3593 from the U.S.; +81 (0)3 3492 6949 from Japan; or + 46 (0)18 16 50 00 from Europe and the rest of the world. Or visit us on the Internet: <http://www.biotech.pharmacia.se>.



Uppsala, Sweden. (And the rest of the world)

READER ENQUIRY NO. 471



The p53 gene from 316 breast cancer patients was sequenced using ALF automated sequencing technology. (Bergh J., Norberg, T., Sjogren, S., Lindgren A., Holmberg, L. "Complete Sequencing of the p53 Gene..." *Nature Medicine* 1995; 10:1029-1034.)



The best athletes are better

SIR — The pre-Olympic Briefing¹ was fascinating reading. Indeed, it gave a good insight into the problems and promises of training for elite sports participation and success. But I found some of the statements in "Science tracks down the training dangers" unjustified. It is true that the media portrays elite sports and intensive training in young athletes as potentially dangerous. Given the relative immaturity of a young athlete's skeletal system, bones are at risk of injuries that could result in long-term deleterious consequences². The irresponsibly high level of training and the will of parents to have an elite sporting child are especially dangerous in the production of both acute and overuse injuries. To clarify the effects of intensive training on young athletes, nearly a decade ago the Sports Council financed a mixed longitudinal study to examine the issue. The study, named TOYA (Training of Young Athletes), recruited a sample of 453 children already at the top for their age groups (9 to 17 years of age at the beginning of the study) in their respective sports (football, gymnastics, swimming and tennis) and followed them for three years.

The TOYA team, based at the Institute of Child Health in London, and of which I was the Medical Officer, was expecting a significant number of injuries in these children. The surprising finding was that in this cohort of elite, intensively trained young athletes in physically demanding sports, each athlete was injured on average only once a year for the whole duration of the study^{3,4}. There were no significant differences to be ascribed to chronological or biological age, velocity of growth, or intensity or length of training⁴. On the other hand, club players of the same age trying to reach the elite status (but not enrolled in the TOYA study) were experiencing serious injuries, requiring them to abandon their sport at a relatively early age⁵.

The TOYA team believes that, while the risk of injury is present in any sport, especially if practised at high level, the elite children entering the study were 'musculo-skeletally' gifted, and their musculo-skeletal system, either through training or genetically, is more able to withstand the stresses of the intensive training needed to reach the elite status and to stay there. The less musculo-skeletally gifted children and their coaches, trying, by emulation of the same training routines, to reach the 'elite'

status, impose on their musculo-skeletal system stresses which, eventually, result in overuse injuries.

In the end, high-level training is not good for some, but some children are able to withstand the rigours that top-level sport entails. At present, however, there is no secure way of identifying children whose musculo-skeletal system can stand intensive training, and we have to learn how much is enough from observing how much is too much.

Nicola Maffulli

Department of Orthopaedic Surgery,
University of Aberdeen Medical School,
Polwarth Building, Foresterhill,
Aberdeen AB9 2ZD, UK

Greek needs

SIR — Current thinking on the future research targets of the European Union (EU) apparently revolves around 'top-down' directives (see *Nature* 381, 631, 634 & 635; 1996). Although this approach is essential for the management of mega-projects with technically visible goals (such as the yeast genome project or high-definition television), it can lead in the long term to asphyxiation of innovative drive. What data support the notion that 'proxy customers' can have a clear perception of the 'products' they will need 10 years down the line? Who among the people in Brussels can predict which plan or bacterial biochemical will produce a future antibiotic? It was not the wishes and visions of end-users or think-tanks that brought about the personal computer revolution, the personal stereo or the World Wide Web. The US biotechnology industry did not blossom because of federal government inspiration.

Wealthier EU members will be well advised to continue supporting fundamental research. Forcing European research to become more 'applied' and 'product-oriented' seems more a cure for the symptoms than for the underlying causes of poor industrial innovation. Of course EU academics must become more aware of the marketplace and develop entrepreneurial thinking and links to corporate Europe. Admittedly, a lot of time and money is lost preparing all these EU collaborative grant applications, but Brussels should seriously consider the introduction of informatics tools to aid submission, refereeing and assessment of research applications. The scientific community is already using the Internet to put proposals together in electronic form before sending printouts to Brussels.

Discussion of the 'cohesion' aspects of the new EU Framework programme should not stem from the assumption that

the inhabitants of less developed EU regions have a genetic predisposition to poor academic performance. This idea will not go down well with the hundreds of young Greeks, Spaniards and Portuguese who are burning the midnight oil in many of the centres of excellence of the North. What the EU should do, instead, is to support the most promising among these young scientists upon their repatriation with 'return grants' not for 12 months (as now) but for five years. Such grants could be administered by the European Molecular Biology Organization, which has gained valuable experience from its outstanding postdoctoral programme.

EU financial support has been instrumental in catalysing and maintaining new research in the 'less favoured' regions of the EU. Establishments such as the Institute of Molecular Biology and Biotechnology in Crete have demonstrated a remarkable ability to carry out research of international calibre and to earn competitive grants, and have trained graduate students and postdoctoral fellows from other EU countries. Such centres have now started attracting reinforcements from the vast pool of expatriate Greek scientists (approximately 6,000 PhDs in the United States; 300 university professors in the Boston area alone). Without EU help, the brain-drain to the United States (where, incidentally, California does not team up with Massachusetts to form elite clubs seeking dominance over Montana) is bound to continue.

Young Greek scientists do not seek charity. We do, however, ask for assistance to grow. Despite admirable efforts by the Hellenic Research Secretariat, we find it hard to get assistance from a government that is forced to maintain the highest defence spending among NATO (North Atlantic Treaty Organization) countries (4.6 per cent of gross national product).

We believe that EU funds should nourish only scientific excellence, overriding the usual politicking that is rife throughout the Old World (and not just on its periphery). And we, in a 'less favoured' region, can carry out good research and achieve technological breakthroughs. Should we be penalized for being successful? Young Greeks, like other young Europeans, have been brought up to believe in the social engineering vision of a community of European states. We are determined to be part of our common tomorrow and to help shape it. We have a lot to offer European science besides memorable holidays.

Tassos Economou

Institute of Molecular Biology
and Biotechnology,
Foundation of Research
and Technology-Hellas,
PO Box 1471, Iraklio,
Crete, Greece
e-mail: aeconomou@nefeli.imbb.forth.gr

1. *Nature* 382, 12–16 (1996).

2. Maffulli, N. *Sports Med.* 9, 229–243 (1990).

3. Baxler-Jones, A., Maffulli, N. & Helms, P. *Arch. Dis. Child.* 68, 130–132 (1993).

4. Maffulli, N., King, J. B. & Helms, P. *Br. J. Sports Med.* 28, 123–136 (1994).

5. Maffulli, N., Chan, D. & Aldridge, M. J. *J. Pediatr. Orthopaed.* 12, 344–350 (1992).

Under arrest at atomic resolution

David O. Morgan

Cell proliferation in many tissues is regulated by p27 protein, which binds and inhibits cyclin-dependent kinases. A crystal structure reveals the remarkable mechanisms by which p27 paralyses these enzymes, arresting the cell cycle.

ENZYMES are plastic. Those elegant helices, strands and loops in crystallographic snapshots are not as tightly knit as they seem, and if the right regulatory subunit comes along, they can undergo dramatic rearrangements. This point has been illustrated beautifully in past studies of the protein kinase Cdk2, and the latest instalment in the Cdk2 story is no exception: on page 325 of this issue¹, Pavletich and colleagues treat us to the structure of Cdk2 in a complex with cyclinA and the inhibitory subunit p27. The structure reveals that p27 is not merely a competitive inhibitor of enzyme activity, but a thoroughly disruptive force that folds into and rearranges the enzyme core.

Cdk2 is one of the cyclin-dependent kinases (CDKs), a family of enzymes whose members govern the transitions of the eukaryotic cell cycle — from the replication of DNA to the processes of cell division. To ensure the proper timing and coordination of cell-cycle events, CDK activity is extensively regulated by a variety of mechanisms². In its monomeric state, the CDK catalytic subunit is inactive. Partial activation is achieved by association with a cyclin; complete activation requires phosphorylation, by a separate kinase, on a conserved threonine residue (Thr 160 in Cdk2). The active CDK–cyclin complex can then be turned off by several mechanisms, including association with CDK inhibitory proteins³, which have achieved considerable notoriety as tumour suppressors and mediators of anti-mitogenic signals that arrest progress through the cell cycle.

The inhibitor p27, whose structure is now reported, was identified just over two years ago as a Cdk2 inhibitor in cells treated with the anti-mitogenic factor TGF- β (transforming growth factor- β), and has since been implicated in a number of anti-mitogenic pathways. Its physiological importance was confirmed two months ago when several groups reported that disruption of the p27 gene in mice leads to an increase in cell proliferation in many tissues^{4–6}.

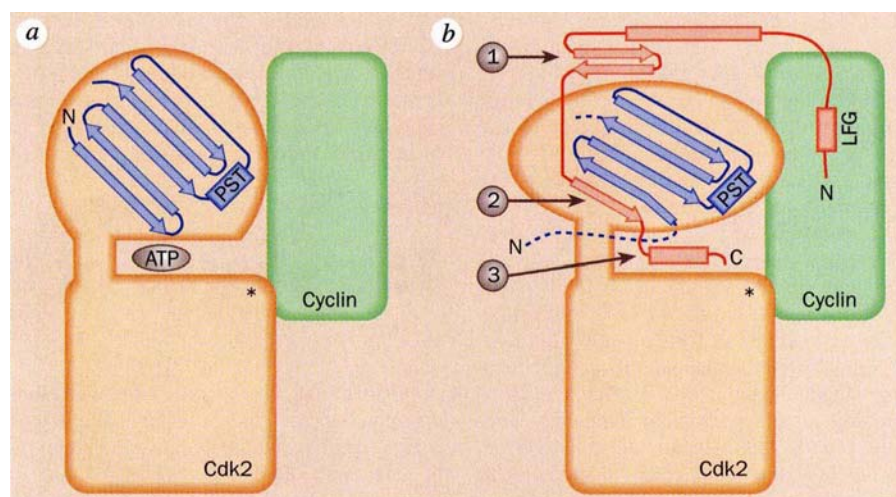
The diverse mechanisms that control Cdk2 activity make this enzyme a prime target for structural studies. Due largely to the efforts of Pavletich and his colleagues, we have recently witnessed explosive progress in this area. We now have structures of Cdk2 in four distinct states

of activity: the inactive Cdk2 monomer⁷, the partially active Cdk2–cyclinA complex⁸, the fully active Thr 160-phosphorylated Cdk2–cyclinA complex (reported in the August issue of *Nature Structural Biology*⁹), and the inactive Cdk2–cyclinA–p27 complex¹. Together these structures provide a comprehensive view of the striking conformational flexibility underlying the regulation of Cdk2.

Like other protein kinases, Cdk2 con-

These problems are almost completely reversed by binding of cyclinA, which reorients the PSTAIR helix and adjoining residues into an active conformation and forces the T-loop down into a relatively unobtrusive position at the base of the active-site cleft⁸. CyclinA is the inflexible partner in this dance: its structure is unaffected by binding of Cdk2¹⁰.

Although cyclin binding alone leads to extensive activation of Cdk2, phosphory-



The effects of p27 on the structure of the Thr-160-phosphorylated Cdk2–cyclinA complex. *a*, A highly simplified view of the active, Thr-160-phosphorylated complex⁹. Cdk2 (yellow) contains two lobes separated by the active-site cleft; the upper amino-terminal lobe contains a five-stranded β -sheet (blue) and the PSTAIR helix ('PST'). Phosphorylated Thr 160 (labelled with a large asterisk) is found in the T-loop at the base of the cleft, where protein substrate presumably binds. The p27-bound complex¹ is shown in *b* (p27 is in red). The LFG motif near the amino terminus of the p27 peptide binds cyclin. In the carboxy-terminal half of the p27 peptide, the three substructures that interact with Cdk2 are numbered as follows. (1) The β -turn of p27 forms a sandwich with the Cdk2 β -sheet. (2) The β -strand of p27 displaces the first strand of the Cdk2 sheet, which is now disordered (dashed line). (3) The p27 3_{10} helix occupies the ATP-binding site.

tains a small amino-terminal lobe drawn by convention above a larger carboxy-terminal lobe. The cleft between the two lobes contains the enzyme's active site, which catalyses transfer of a phosphate from ATP, bound deep within the cleft, to the protein substrate. The structure of the Cdk2 monomer unveiled two major impediments to activity⁷. First, a helix in the small lobe above the active-site cleft — the 'PSTAIR helix' (in single-letter amino-acid code) — is twisted out of position, resulting in the misorientation of side chains involved in ATP binding. Second, a long loop (the 'T-loop') rises from the carboxy-terminal lobe to prevent protein substrate binding at the entrance of the active site.

lation of Thr 160 is required for full activation *in vitro* and complete CDK function *in vivo*. The new structure of the phosphorylated Cdk2–cyclinA complex tells us why⁹. The threonine at position 160 is located in the T-loop at the base of the active-site cleft; the phosphate at this site fits snugly into a positively charged pocket formed in part during cyclin binding, dragging the T-loop farther down into a position where it no longer blocks the protein substrate-binding site. A network of hydrogen bonds fans out from the phosphate to stiffen the surrounding region of Cdk2 and stabilize its association with cyclinA.

Now on to the latest chapter in the Cdk2 saga: the effects of the inhibitor p27

on the Thr-160-phosphorylated complex between Cdk2 and cyclinA¹. The p27 structure in this work is limited to a 69-amino-acid peptide known to be sufficient for Cdk2 inhibition. When bound to Cdk2–cyclinA, the p27 peptide is draped across the tops of the two subunits in an extended conformation (see the accompanying figure, and Fig. 2a on page 327). The amino-terminal region of the p27 peptide binds cyclinA (which, once again, remains unmoved by the experience). A key component of this interaction is a short amino-acid motif on p27 that is conserved in related inhibitor proteins; the central sequence in this motif ('LFG') is inserted in a binding pocket on the face of cyclinA. Interestingly, the LFG motif is found in other proteins that bind cyclinA, including the retinoblastoma-related proteins p107 and p130 (ref. 11). These proteins are thought to be CDK substrates, and presumably the association with cyclin enhances their phosphorylation; the binding of p27 to the same site on cyclin would therefore result in competitive inhibition of protein substrate binding.

While the amino-terminal region of the p27 peptide anchors the inhibitor to the cyclin, the carboxy-terminal region wreaks havoc with the small upper lobe of the Cdk2 subunit (see figure). First, the 'β-turn' segment of p27 flattens the β-sheet atop the kinase. Second, the 'β-strand' segment of p27 shoves aside the first β-strand in the Cdk2 sheet, which now flaps disordered in the solvent. These injuries alone severely destabilize ATP binding. To make matters worse, the third carboxy-terminal segment of the p27 peptide, the '3₁₀ helix', binds deep within the active-site cleft to dash any hope of ATP binding (see Fig. 5b, page 330, for a striking illustration). So a short stretch of amino acids in p27 employs every imaginable trick to shut down kinase activity.

The extensive interactions between p27 and both subunits of the complex are consistent with observations that the inhibitor binds more avidly to the complex than to either subunit alone^{12,13}. It is also consistent with evidence that these inhibitors can, under some conditions, enhance the binding of Cdk2 and cyclinA (refs 12, 13). Does this mean that p27-like inhibitors are CDK–cyclin 'assembly factors' in the cell? This seems unlikely in the case

of Cdk2–cyclinA complexes, which are known to bind with very high affinity in the absence of other factors, but inhibitor-induced assembly might be important for some CDK–cyclin pairs whose intrinsic affinity is not so high.

The structure of the complex between Cdk2, cyclinA and p27 was anxiously awaited by those of us who hoped that it would clear up some of the confusion about the stoichiometry of inhibitor–CDK binding. Under certain conditions, it is possible to isolate inhibitor–Cdk2 complexes that possess kinase activity^{12,13}. Complete inhibition of these complexes seems to require the binding of two inhibitor molecules. The new structure, however, provides striking evidence that a single p27 can completely inhibit Cdk2. Can we explain this discrepancy?

It is conceivable that p27 does not always reach the fully bound state: perhaps under some conditions the inhibitor is bound only to the non-inhibitory binding site on the cyclin, in which case a second p27 could interact with the inhibitory site on the CDK. Clearly, the binding of a different p27 molecule to each of the two binding sites is thermodynamically far less favourable than the occupation of both

sites by a single p27. It remains possible, however, that partial association of p27 with the CDK, perhaps enhanced by certain experimental conditions, could yield a partially active inhibitor–Cdk2 complex.

With four Cdk2 structures completed, it is now clear that regulatory subunits and modifications can induce striking changes in the shape of the CDK active site. What comes next? If we hope to achieve a complete understanding of CDK regulation, then there are at least two major CDK inhibitory mechanisms that remain unexplored at this resolution. To begin with, we do not understand how CDKs are inhibited by phosphorylation at the threonine and tyrosine at positions 14 and 15, the side chains of which dangle near the ATP-binding site. Similarly, little is known about the inhibitory mechanisms employed by the INK4 class of protein inhibitors, which are unrelated to p27 and bind only to the catalytic subunits of certain CDKs. These and related issues promise to keep the X-ray machines running for years to come. □

David O. Morgan is in the Department of Physiology, University of California, San Francisco, California 94143-0444, USA.

COSMOLOGY

The shards of broken symmetry

Wojciech H. Zurek

A COSMOLOGICAL experiment sounds like an impossibility. Yet precisely such a feat is described in two complementary papers in this issue. Two groups, one based in Grenoble¹ (page 332), the other in Helsinki² (page 334), report on a pair of superfluid helium-3 experiments, devised to probe the dynamics of the cosmological phase transitions that are thought to have occurred a fraction of a second after the Big Bang.

As the Universe expanded and cooled, the fields that mediate fundamental forces (believed to be unified at a very high temperature) underwent a sequence of symmetry-breaking phase transformations. The effect was to select the particular broken-symmetry vacuum of the Universe we inhabit, with its gravitational, electromagnetic, weak and strong interactions, and their associated entourages of elementary particles. As Tom Kibble pointed out³, phase transitions can lead to topological defects (monopoles, cosmic strings or domain walls) which can have either desirable or devastating consequences for different cosmological theories.

Spontaneous symmetry breaking is familiar in the context of condensed matter. A ferromagnet cooled below the critical temperature is the classic example. As thermal fluctuations decrease, it becomes

energetically favourable for different elementary magnets to point in the same direction, selected at random; the choice of direction breaks the spherical symmetry

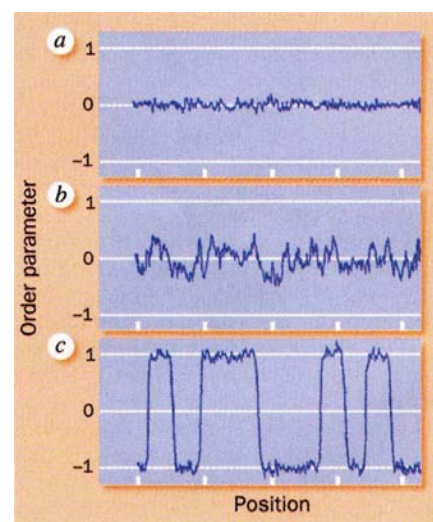


FIG. 1 Numerical simulation of the formation of topological defects (known as 'kinks' in this one-dimensional case) during a phase transition¹⁰. a, The order parameter ψ at high temperature oscillates around zero. b, As the transition gradually takes place, ψ must choose between +1 and -1. c, When domains with different values of ψ meet, a kink connects them.

1. Russo, A. A., Jeffrey, P. D., Patten, A. K., Massagué, J. & Pavletich, N. P. *Nature* **382**, 325–331 (1996).
2. Morgan, D. O. *Nature* **374**, 131–134 (1995).
3. Sherr, C. J. & Roberts, J. M. *Genes Dev.* **9**, 1149–1163 (1995).
4. Nakayama, K. et al. *Cell* **85**, 707–720 (1996).
5. Kiyokawa, H. et al. *Cell* **85**, 721–732 (1996).
6. Fero, M. L. et al. *Cell* **85**, 733–744 (1996).
7. De Bondt, H. L. et al. *Nature* **363**, 595–602 (1993).
8. Jeffrey, P. D. et al. *Nature* **376**, 313–320 (1995).
9. Russo, A. A., Jeffrey, P. D. & Pavletich, N. P. *Nature Struct. Biol.* **3**, 696–700 (1996).
10. Brown, N. R. et al. *Structure* **3**, 1235–1247 (1996).
11. Zhu, L. et al. *Genes Dev.* **9**, 1740–1752 (1995).
12. Zhang, H., Hannon, G. J. & Beach, D. *Genes Dev.* **8**, 1750–1758 (1994).
13. Harper, J. W. et al. *Mol. Biol. Cell* **6**, 387–400 (1995).

that the ferromagnet has above the critical point. But theories that govern phase transitions in cosmology and in condensed-matter physics are essentially identical, so experiments on cosmology can in theory be performed in a laboratory.

Common to cosmology and condensed-matter physics is the question, how big are the pieces of broken-symmetry vacuum just after the phase transition? The size of such coherent domains manifests itself in the initial density of topological defects. These are intrusions of the old symmetric vacuum, imprisoned by configurations of the new broken-symmetry state that are impossible to disentangle. For example, when symmetry breaking involves two distinct choices, at a boundary between the two such regions there will be a domain wall (Fig. 1). One can expect approximately one 'unit' of topological defect — one monopole, a domain-sized piece of a string, or a 'tile' of a domain wall — per domain³, so the initial density of defects gives the size of the pieces of broken symmetry.

Superfluids are an excellent example of symmetry breaking that leads to topological defects. They are described by a complex 'order parameter', $\psi = |\psi|e^{i\theta}$, a wave function of the Bose condensate of helium atoms. In some ways, ψ can be thought of simply as a probability wave familiar from ordinary quantum mechanics. The square of its absolute value gives the density of the superfluid, which in a sense is the probability of finding an atom of superfluid at a particular point. But the analogy is incomplete. The wavefunction of the

superfluid Bose condensate does not satisfy the superposition principle, and its square is not conserved. Moreover, in helium-3 the order parameter ψ is not just a complex number, but a tensor, which describes rotational states of the pairs of ³He atoms (individual ³He atoms are fermions, and like electrons in a superconductor they must pair before they can Bose-condense).

Symmetry breaking is induced by a change in the shape of the effective potential. Above the critical temperature, the potential minimum is at the origin and ψ fluctuates near there (Fig. 2), but below the critical point the potential has a minimum away from the origin, along a rim of the 'sombrero'. To settle there, ψ has to break symmetry — it has to choose a value of θ .

Topological defects in superfluids arise when the phase, θ , changes continuously by 2π along a closed path in space — then there must be a point inside the loop where θ is undefined. The arrows representing the order parameter will all be pointing either directly towards or away from the 'vortex

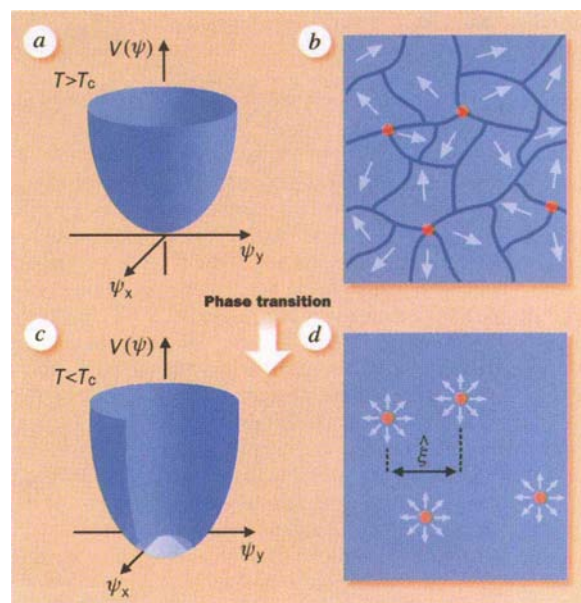


FIG. 2 Dynamics of defect formation in superfluids. *a, b*, Above the phase transition, the two-dimensional (complex) order parameter ψ fluctuates around the minimum of the potential, assuming values which point in approximately the same direction (or phase, θ) within each local domain. *c, d*, When this configuration is frozen out by the phase transition, the symmetric phase will be trapped in some locations: these are defects. In superfluids, such defects are lines of vortical flow. Their energy is related to the difference in potential energies between the hill at $\psi = 0$ and the minimum along the circle of broken-symmetry configurations. It can also be regarded as the energy of rotation of the vortex.

line', where the field is forced to climb back on top of the sombrero to assume the only value, $\psi = 0$, that does not require making a choice of phase. A gradient in θ necessarily implies a flow, just as an ordinary plane wave flows in the direction of changing phase. In this case the flow pattern is a vortex of superfluid with a non-superfluid core.

The energy corresponding to this rotation was used by the Grenoble group¹ to detect the presence of vortex lines — superfluid analogues of cosmic strings. This can be done by heating up a small volume of superfluid ³He and letting it cool down rapidly, so for a short while the system is in a non-superfluid state. When it returns below the critical temperature, different pieces independently choose the manner in which their symmetry is spontaneously broken. The initial energy used to heat up the fragment of the superfluid above the critical temperature is precisely known, so by monitoring the energy that escapes, one can find out how much of it remains in the vortex lines. The Helsinki group uses an identical procedure to generate defects², but they do it at a somewhat higher temperature, and they use a technique that allows them to *count* the loops of string-like vortex lines one by one. In both experiments the initial density of vortex lines is found and used to infer the average size of the like-minded pieces of newly made broken-symmetry superfluid.

The answer confirms a theory of the

To derive the density of defects:

Like many of the fundamental fields in cosmology, superfluids undergo what are called second-order phase transitions. As such a system approaches a transition infinitesimally slowly, its order parameter becomes correlated over larger and larger domains. The domain size is characterized by the correlation length ξ , which, in equilibrium, increases and diverges to infinity as the temperature approaches the critical temperature T_c . More precisely, in ³He and many other systems $\xi = \xi_0 \epsilon^{-1/2}$, where ξ_0 is a constant and ϵ is the relative temperature, $\epsilon = (T - T_c)/T_c$. So, if the temperature changes slowly enough through T_c , ξ diverges to infinity, and in any finite volume there will be no defects. However, it also takes an increasingly long relaxation time τ for the system to settle to such a uniform equilibrium state, as $\tau = \tau_0/\epsilon$. Here, τ_0 is the characteristic reaction time far below the phase transition.

If the transition occurs at a finite pace, on a quench timescale τ_Q , then $\epsilon = t/\tau_Q$, where t is time. There will

come a moment $\hat{t}$ when the system will not be able to keep up with the changes in ϵ — its reaction time will simply be too slow. That will happen when the time remaining until the phase transition is equal to its reaction time, $\tau(\hat{t}) = \hat{t}$ or $\hat{t} = \tau_0 \tau_Q / \hat{t}$. Hence, $\hat{t} = (\tau_0 \tau_Q)^{1/2}$. Before the critical temperature is reached, the state of the system will fall out of step with what the equilibrium should be, and by the time its reflexes are recovered it will be already in the broken-symmetry state. The correlation length corresponding to $\hat{t}$ should then be the size of the domains⁴, the 'effective causal horizon' for the order parameter. That is, the size of the uniform domain $\hat{\xi} = \xi_0 \epsilon^{-1/2} = \xi_0 (\tau_Q / \tau_0)^{1/4}$ defines the extent of the region within which the order parameter has managed to agree on how to break the symmetry. Vortex lines in ³He should form with a density of about one $\hat{\xi}^3$ of length per $\hat{\xi}^3$ volume. This is consistent with the numerical simulation¹⁰ and both the new experiments^{1,2}. W.H.Z.

initial domain size based on a combination of causality and equilibrium thermodynamics⁴ (see box). A rough confirmation of these ideas from ⁴He experiments was already at hand⁵, but ⁴He has a very different set of experimental parameters, which allow for generation of a much more plentiful vortex tangle but which also make the subsequent determination of their density more difficult. Thus, it was only possible to set a lower limit on the initial density of vortex lines.

The three experiments^{1,2,5} together confirm the simple ideas outlined in the box^{3,4,6}, and open up a new area of inquiry: the non-equilibrium aspects of phase transformations. Their implications go beyond the issue of topological defects. For instance, it may be possible to restore the chiral symmetry (corresponding to the quark-hadron transition) of a nucleus-sized piece of the Universe in accelerator experiments in the near future⁷. We could then see chiral symmetry breaking of the vacuum, resembling phenomena in the heated bubble of ³He. The analogy is not complete — in this case the sombrero potential is tipped and there is no possibility of defects, but this was one of the last transitions to occur as the Universe cooled. Moreover, the size of the pieces of the broken-symmetry chiral vacuum can translate into measurable fluctuations in the number of pions.

The experimental confirmation of the mechanism for defect formation alone does not determine whether they exist in our Universe; for that, one needs to know more about the Grand Unified Theory and its relationship to the present-day low-temperature vacuum^{8,9} — it is the topology of the broken-symmetry vacua that decides whether the order parameter can disentangle itself, or whether it has to leave behind traces of discord, crystallized out as, say, cosmic strings. But we can now be confident that when defects are possible in principle, they will appear with an initial density that can be predicted by a theory now tried and tested in superfluids. □

Wojciech H. Zurek is in the Theoretical Astrophysics Group, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA.

Endless forms, several powers

Chung-I Wu and Norman A. Johnson

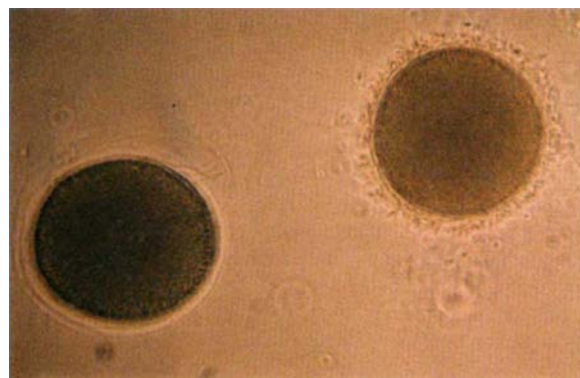
NEARLY all biologists agree that speciation matters — after all, this is how the diversity of life comes about. Beyond that, however, there is plenty to argue about, a recent forum being a meeting* held in California in honour of Guy L. Bush.

Historically, the debates have been especially intense on three fronts. First, is speciation at all possible in the presence of gene flow between co-existing species-in-the-making (sympatry)? It has never been in doubt that speciation often takes place when two populations become geographically separated (allopatry). But how might sympatric speciation actually happen in those cases that can be studied?

species of the true fruitfly, *Rhagoletis*, evolved in sympatry through a process involving rapid shifts of host; they further suggest that the apple-host race of *R. pomonella* probably evolved only in the last two centuries. There is no lack of scepticism about the claims. For example, there may be microgeographical barriers to gene flow, meaning that speciation may not be truly sympatric. On the other hand, if there is gene migration between the species, the so-called races could be no more than polymorphisms of the same species.

New evidence suggests that genetic differentiation has indeed taken place despite substantial gene flow between races of up to 6% per generation (J. Feder, Univ. Notre Dame), corroborating earlier interpretations^{1,2}. An effective barrier against the migrant genes seems to be the differential timing of fruit maturation and, consequently, the timing of emergence of the fly. Significantly, this host-shift process may have been played and replayed several times in the *Rhagoletis* species complex (S. Berlocher, Univ. Illinois).

Sympatric speciation by host shift is not a new or radical idea (it goes back to Charles Darwin). But what of other lines of evidence



Left out — the evolution of gamete-recognition proteins may contribute to speciation in coexisting Hawaiian sea urchins. The egg on the right is subject to the frenzied attentions of sperm from the same species. The egg on the left, from a sister species, has not been recognized. Such gamete incompatibility is one plausible cause of sympatric speciation.

Second, how many genes does it take to make a new species? Are nascent species really not very different, with but a small number of crucial changes separating them, or are they products of extensive genetic evolution? Third, what are the evolutionary forces that drive species apart? And how does the selection of and competition for mates (sexual selection) drive the process? These are, in fact, highly interrelated questions. For example, sympatric speciation would be more plausible if speciation requires only a small number of genes than if extensive genetic changes are necessary.

In the past, discussion of these issues has been characterized as much by speculation as by evidence. So it was refreshing to find at the meeting that, increasingly, empirical data as well as theoretical considerations are being brought to bear on the debates.

Bush (Michigan State Univ.) and his associates have long maintained that sibling

that sympatric speciation occurs in other circumstances? There are many cases of rapid speciation in aquatic environments where no apparent barriers to gene flow exist, as between two sister species of cichlids in a crater lake (ref. 3) or among the 26 endemic species of sculpins in Lake Baikal (A. McCune, Cornell Univ.). Another example is between highly migratory species of marine invertebrates (S. R. Palumbi, Univ. Hawaii).

What is unique about marine invertebrates is the simple genetics of gamete recognition, for instance by the sperm acrosomal protein, bindin, in sea urchins. DNA sequences of bindin show that the species recognition signal is driven by strong selection, and different alleles confer different affinities with eggs of different females (see figure). Sympatric speciation is theoretically feasible if the right genetic assumptions can be made (A. Kondrashov, Cornell Univ.; P. Johnson, Michigan State Univ.). Palumbi cited a model that would indeed allow sympatric speciation under the sort of genetic basis he observed⁴: the model requires

1. Bäuerle, C., Bunkov, Yu. M., Fisher, S. N., Godfrin, H. & Pickett, G. R. *Nature* **382**, 332–334 (1996).
2. Ruutu, V. M. H. et al. *Nature* **382**, 334–336 (1996).
3. Kibble, T. W. B. *J. Phys. A* **9**, 1387–1389 (1976).
4. Zurek, W. H. *Nature* **317**, 505–508 (1985).
5. Hendry, P. C., Lawson, N. S., Lee, R. A. M., McClintock, P. V. E. & Williams, C. H. D. *Nature* **368**, 315–317 (1994).
6. Zurek, W. H. *Phys. Rep.* (in the press). Preprint cond-mat/9607135 at xxx.lanl.gov
7. Bjorken, J. D. *Acta Phys. Pol. B* **23**, 561–572 (1992).
8. Vilenkin, A. & Shellard, E. P. S. *Cosmic Strings and Other Topological Defects* (Cambridge Univ. Press, 1994).
9. Kibble, T. W. B. in *Formation and Interaction of Topological Defects* (eds Davis, A. C. & Brandenberger, R.) 1–26 (Plenum, New York, 1995).
10. Laguna, P. & Zurek, W. H. *Los Alamos Preprint LAUR 96-2036* (1996). Preprint gr-qc/9607041 at xxx.lanl.gov

*Endless Forms: Species and Speciation, Asilomar, California, USA, 19–23 May 1996.

multiple allelic determinants and large population size, both of which are fulfilled in this case.

The genetics of divergence between closely related species is, naturally, a constant theme. One issue is to do with the role of the X chromosome, as opposed to the autosomes, in speciation (D. Prowell, Louisiana State Univ.; M. Ritchie, Univ. St Andrews). The X does not play a disproportionate part in pre-mating reproductive isolation, but it was widely thought that it does in post-mating isolation. New studies^{5,6} showing comparable densities of hybrid sterility factors on all chromosomes run counter to that view.

A central question is whether species differences, especially traits that are responsible for reproductive isolation, are due to a relatively small number of genes that have strong effects or a large number of genes that have weak effects. Three high-resolution studies have reached very different conclusions. In one, each of the eight traits differentiating two species of monkey flowers has a locus that determines more than 25% of the difference (D. Schemske, Univ. Washington). On the other hand, reduction of hybrid fertility between two *Drosophila* species may follow a gradation model, with at least six loci needed to reduce male fertility noticeably (H. Naveira, Univ. La Coruña, Spain). Likewise, between two other sibling species of *Drosophila* that have been subjected to DNA marker-assisted mapping, the number of loci contributing to hybrid sterility is more than 120; the genetics in this case can be described as 'weak-allele and strong-interaction' (C.-I. Wu).

The contrasting conclusions could be due to a number of factors, most intriguing among them being the difference between plants and animals. The creation of new species by hybridization may be quite common in plants, as their hybrids are sometimes better adapted to either the original or novel habitats (M. Arnold, Univ. Georgia). Such instances are much rarer in animals, implying very different genetic architectures for species divergence. The difference may also be part of the reason that botanists tend to be more receptive to the possibility of sympatric speciation and do not adhere to the species-isolation definition as closely as do zoologists (A. Templeton, Washington Univ.).

Finally, what of the evolutionary forces driving speciation — sexual selection, in particular? As noted above, coevolution of sexual traits has been suggested to be the driving force of sympatric speciation among sea urchins. In a dramatic demonstration of the power of sexual selection in *Drosophila*, males evolving without the coevolution of females can rapidly become harmful to females (W. Rice, Univ. California, Santa Cruz)⁷. And there are several examples of coevolving male and female traits in desert *Drosophila* (T. Markow, Arizona State Univ.), in that species with less responsive females usually have males with greater ability to stimulate oviposition. This is true with mating latency as well, potentially accounting for the asymmetry in reproductive isolation.

Male-male competition is another component of sexual selection. There is a case, for example, where two species of crickets do not appear to be exchanging genes, and hybrids are produced at a low level. Most intriguing of all, there is no sign of pre-mating isolation and the hybrids are normal. How, then, are the species reproductively isolated? As it turns out, when females are inseminated by many males (females mate constantly), conspecific sperm enjoys a very large

advantage in fertilizing eggs⁸ (D. Howard, Univ. New Mexico). That leaves the question of why, given the strong post-mating isolation, pre-mating isolation has not evolved. There are also instances of extensive genetic divergence underlying sexual isolation between races of *Drosophila*, but without accompanying post-mating isolation, showing that the two phenomena can be somewhat decoupled (they are often linked by the reinforcement argument⁹, according to which pre-mating isolation follows post-mating isolation).

The meeting thus leaves a vivid impression of the diversity of patterns and processes underlying speciation — allopatry or sympatry, simple or complex genetics, sexual or natural selection and so on. The task is to sort out their relative importance and the actual mechanics of each process. Did Darwin not invoke several powers to close the *Origin of Species*? □

Chung-I Wu is in the Department of Ecology and Evolution, University of Chicago, Chicago, Illinois 60637, USA. Norman A. Johnson is in the Department of Biology, University of Texas, Arlington, Texas 76019, USA.

CHEMISTRY

Riding tandems with nature

Helmut Ringsdorf

A NATO workshop¹ on tandem molecular interactions was held at Il Ciocco in Italy last month*. Its purpose was to bring together chemists, physicists and theoreticians — mostly from the field of materials science — to examine the simultaneous use of two or more molecular interactions to adjust polymer organization and thereby create sophisticated polymer systems. A new concept? Certainly not — at least not for mother nature, where synergistic or competitive tandem processes and interactions are the basis for most cell and membrane chemistry. One of the best examples is the acetylcholine receptor². The release of acetylcholine at the presynaptic membrane of nerves induces the direct activation of sodium channels, thereby causing depolarization of the postsynaptic membrane and passing on the signal.

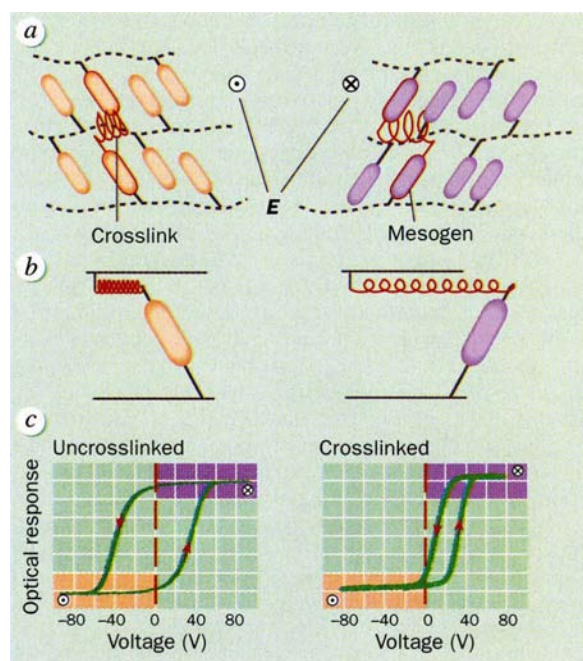
By tandem now to new materials? That question was the basic motivation for this workshop, which developed from discussions between R. Zentel, C. K. Ober and G. Galli in 1994. Polymers can organize on a variety of length scales. An important example is the class of block copolymers that

have liquid-crystal (LC) segments, where the polymer chain has rod-like side groups, called mesogens, which can align in a regular array. Blocks (domains) form with either liquid-crystalline or amorphous packing. The different blocks would be immiscible, but they cannot phase-separate macroscopically because they are covalently linked. So locally, at ten-nanometre scales, this material is either amorphous or in the liquid-crystal phase, but superimposed on this is a regular biphasic structure with domain sizes of several hundred nanometres. Until now such systems have been discussed as isolated topics.

The domains in block copolymers can form a variety of shapes — spheres, rods or layers. They owe their morphology to microphase separation and the tandem, competing interaction of interfacial tension and the entropic forces caused by chain stretching. The first tries to minimize the interface between the microdomains, the second, as the chains coil up and 'fatten', tries to increase it. This interplay between two forces, and the resulting morphologies, can be manipulated chemically, by changing the connectivity between the blocks for example. Polymers in which the main polymer chain (rather than the side groups) forms the liquid-

1. Feder, J. L., Chilcote, C. A. & Bush, G. L. *Nature* **336**, 61–64 (1988).
2. McPherson, B. A., Smith, D. C. & Berlocher, S. H. *Nature* **336**, 64–66 (1988).
3. Schilwen, U. K., Tautz, D. & Paabo, S. *Nature* **368**, 629–632 (1994).
4. Wu, C.-I. *Evolution* **39**, 66–82 (1985).
5. True, J. R., Weir, B. S. & Laurie, C. C. *Genetics* **142**, 819–837 (1996).
6. Hollocher, H. & Wu, C.-I. *Genetics* **143**, 1243–1255 (1996).
7. Rice, W. R. *Nature* **381**, 232–234 (1996).
8. Robinson, T., Johnson, N. A. & Wade, M. J. *Heredity* **73**, 155–159 (1994).
9. Kelly, J. K. & Noor, M. A. F. *Genetics* **143**, 1485–1497 (1996).

*Manipulation of Organization in Polymers Using Tandem Molecular Interactions, Il Ciocco, Tuscany, Italy. 28 May to 2 June 1996.



A ferroelectric LC-elastomer and its two switching states. *a*, A polymer chain acts as crosslinking point by connecting different mesogenic groups, attached to the main polymer chains. For an uncrosslinked polymer (without the additional chain acting as crosslinking point), both switching states are equivalent. For a crosslinked LC-elastomer, however, ferroelectric switching extends polymer chains. *b*, The entropy elasticity arising from this acts like a spring, which stabilizes one state and shows up in measurements of hysteresis (*c*). For the uncrosslinked system (left), the hysteresis is symmetric to zero voltage and both states are stable, with no electric field. After crosslinking in one polar state (right), only that state is stable with no electric field, and the hysteresis is no longer symmetric to zero voltage.

crystalline phase are already stretched straight, and can produce new, chevron-like structures which are much stronger under compression than ordinary LC polymers³ (C. K. Ober, Cornell Univ., and E. L. Thomas, MIT).

Traditional block copolymers are used for their exceptional mechanical properties — rubber-toughened polystyrene, for example, is as hard as pure polystyrene, but much tougher, because cracks can only propagate as far as the next rubber domain and so damage is localized. But many other applications are possible. Phase-separated block copolymers with ionic groups can be prepared with very short ionic blocks. Kinetic freezing (that is, cooling to a glassy state) can then be used to make micelles of different shapes (A. Eisenberg, McGill Univ., Montreal)⁴ or to form

lateral structures in thin films⁵ (M. Möller, Univ. Ulm). Either of these can be used as minireactors, to keep reacting species in place to form nanoparticles of cadmium sulphide or gold, for example.

In liquid crystals, the crosslinking of polymer chains transforms a liquid into a soft rubbery solid, while still allowing enough mobility for liquid-crystal phase transitions. The elasticity of such three-dimensional networks should stabilize the 'smectic A phase', in which there is one-dimensional long-range order perpendicular to the layers, with liquid-like order in the other two dimensions (E. Terentjev, Univ. Cambridge). This LC rubber possesses, therefore, highly directional mechanical properties, combining entropy (rubber-like) and energy (metallic) elasticity in one material (H. Finkelmann, Univ. Freiburg).

In combination with ferroelectric liquid crystals (which can settle into a state that generates a macroscopic electric field), ferroelectric LC-elastomers are possible⁶ (see figure). In these systems, different tandem interactions are active. An external electric field can interact with the ferroelectric liquid crystal and lead to switching of the electric polarization (part *a* in the figure). After crosslinking,

ferroelectric switching leads to a deformation of the polymer network, equivalent to stretching a spring (*b*). This can be seen by comparing the ferroelectric hysteresis before and after crosslinking (*c*).

One possible application of these materials is in the field of rubbery piezosensors⁶ (R. Zentel, Univ. Mainz). As rubber can deform by 100% or more, compared with less than 1% for solid piezoelectric crystals, this is especially promising for electromechanical transducers. They use the inverse piezoelectric effect, in which an applied electric field causes a deformation.

The most provocative contribution, "Can tandem interactions give rise to motion?" (J. Prost, Inst. Curie, Paris), was a theoretical one. It was stimulated by examples from nature like the motor protein assemblies that drive muscles. The hope is to realize motion in artificial systems like microlattices⁷, where tandem interactions can stabilize two different spatial structures at similar energies, allowing them to be chemically cycled back and forth in some sort of molecular motor. But just what the details of a system would be is still far from clear.

Often, when an idea is put into words its importance becomes clear, and aspects of a field become evident that had been overlooked or discussed in a 'phase-separated' manner before. The concept of tandem molecular interactions, which are ubiquitous in nature, has the potential to provide new theoretical perspectives in materials science, as well as materials with new properties. □

Helmut Ringsdorf is at the Institut für Organische Chemie, Universität Mainz, Becher-Weg 18–20, D-55099 Mainz, Germany.

The Earth opens up

THIS imposing statuary greets the visitor to the Earth Galleries at the Natural History Museum in London, which opened last Saturday, 20 July. It stands within the visually arresting atrium, its 15-metre slate walls decorated with stars and planets. An escalator whisks the visitor through an 11-metre-diameter rotating globe



to two new exhibitions: *The Power Within*, on the Earth's interior, volcanoes and earthquakes; and *Restless Surface*, on the forces that shape the landscape. Four further exhibitions will have opened by the end of 1998.

Mounting the exhibitions required the complete refurbishment, at a cost of £12.1 million, of the former Geological Museum, which the Natural History Museum took over from the British Geological Survey in 1985. H.G.

- Höcker, H. (ed.) *Manipulation of Organization in Polymers using Tandem Molecular Interactions*, *Macromol. Symp.* (in the press).
- Sakmann, B., Patlak, J. & Neher, E. *Nature* **286**, 71–73 (1980).
- Chen, J. T., Thomas, E. L., Ober, C. K. & Hwang, S. S. *Macromolecules* **28**, 1688–1697 (1995).
- Zhang, L., Yu, K. & Eisenberg, A. *Science* **272**, 1777–1779 (1996).
- Spatz, J. P., Roescher, A. & Möller, M. *Adv. Mater.* **8**, 337–340 (1996).
- Zentel, R. & Brehmer, M. *Chemtech* **25**(5), 41–48 (1995).
- Rousselet, J., Salome, L., Ajdari, A. & Prost, J. *Nature* **370**, 446–448 (1994).

Simple worms, complex genes

Scott W. Emmons

CLASSICAL results in experimental embryology established long ago that cells of the developing animal have a regional identity. They can be characterized not only as 'skin', 'nerve' and 'bone', but also as 'arm' and 'leg'. But how cells know what body region they belong to, and what to do there, is not known. Results reported in this issue¹ and in *Development*² describe unexpected properties of a key player, one of the *Hox* genes — the dynamic, lineage-based regulation of a *Hox* gene in the nematode *Caenorhabditis elegans* is at odds with a traditional view of *Hox* genes as relatively fixed markers of regional cell identity.

Many genes that act in the specification of cell fate during development have been identified in model organisms amenable to genetic analysis. One such is *C. elegans*, which has the virtue of being made up of a small number of cells (around 1,000) that arise by a reproducible developmental cell lineage. In spite of its small size and fixed lineages, *C. elegans* has a repertoire of developmental regulatory genes that is remarkably similar to that found in larger animals with less determinate cell lineages. These genes include the *Hox* transcription-factor genes, which pattern the anteroposterior body axis, as well as other structures, such as vertebrate limbs. *Hox* genes encode proteins with the DNA-binding protein motif known as the homeobox, and are organized in conserved genomic clusters. Remarkably, and for unexplained reasons, they are generally arrayed along the chromosome in the same order as they act in the body.

Evidence for *Hox* action in specifying regional identity first came from the fruitfly *Drosophila*, where cell-cell interactions seem to be more important in patterning the tissues than in *C. elegans*. In *Drosophila*, mutations in *Hox* genes result in striking transformations of one body region into another. These transformations affect many diverse body structures, and so

have been interpreted as being the result of spatial transformations of an abstract cellular quality variously termed 'regional identity', 'positional identity' or 'positional value'. But the nature of regional identity, and how *Hox* genes confer it, remains largely unknown.

An economical mechanism would be that each cell expresses a certain *Hox* gene which acts as a marker to record that cell's regional identity. Because there are generally more body regions than *Hox* genes, the genes would have to double up to establish some kind of code. This view, first put forward by Lewis³, has guided much of the thinking about *Hox* ever since⁴.

The strength of models postulating *Hox* genes as fixed cellular markers is that they readily explain how *Hox* mutations, or ectopic expression of *Hox* proteins, produce segmental or regional transformations. But they account less easily for other observations. For instance, when *Hox* expression is examined during development in detail at the cellular level, spatially complex patterns that vary over time are seen⁵. Such patterns suggest that regional identity may not, after all, be a property attributable to single cells. It may instead be an emergent property of a group of cells, consisting of a particular spatial pattern or temporal series of regulatory events affecting *Hox* genes. So the question has remained unresolved. Is regional identity a state of a cell or a process proceeding in a group of cells?

A view of regional identity emphasizing the regulation of *Hox* gene expression is now supported by detailed studies in *C. elegans* by Kenyon and co-workers^{1,2}. They have undertaken a detailed analysis of the regulation of *mab-5*, the *C. elegans* homologue of the *Drosophila Antennapedia* gene, during both embryonic and postembryonic development.

A first surprising finding (page 353 of this issue¹) is that, although *mab-5* is normally expressed in cells of the posterior body region during embryogenesis (Fig. 1), regulation of its expression is controlled by a lineal rather than a positional mechanism. Cowing and Kenyon¹ took advantage of

changes in embryonic blastomere fates that occur after early blastomere ablations to generate *mab-5*-expressing lineages at abnormal positions within partial embryos. They found that *mab-5* was expressed from the appropriate lineage branch regardless of where this occurred. Thus, *mab-5* expression cannot be solely induced by an extrinsic signal specific to the posterior region of the embryo.

These experiments examined cells of the ectoderm. In further work, Cowing

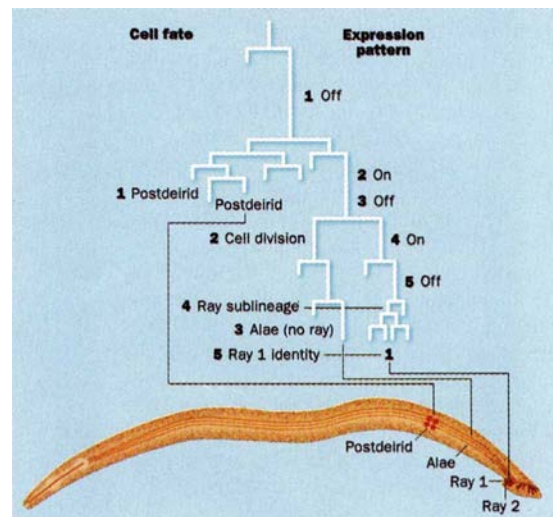


FIG. 2 In the second paper discussed here, Salser and Kenyon² show that, during postembryonic development, *mab-5* expression in one ectodermal cell lineage cycles on and off repeatedly to specify several adult structures as well as a cell division. The postdeirid and rays are neuronal structures; alae are cuticular ridges. At the very end of development, expression of *mab-5* from a heat-shock promoter results in the transformation of the morphological identity of ray 1 to that of ray 2.

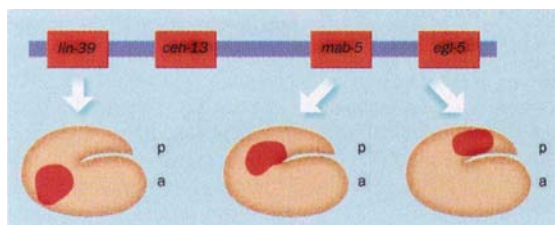


FIG. 1 Contrary to models postulating *Hox* genes as fixed cellular markers of regional identity, a *C. elegans* *Hox* gene is regulated in a complex manner by cell lineage. In *C. elegans*, *Hox* genes are arrayed in a chromosomal cluster as in other animals, and are expressed during embryonic development in specific body regions. Cowing and Kenyon¹ show that *mab-5*-expressing cells, though they are found in one region of the embryo, are not specified by region-specific signals but by a lineage mechanism.

and Kenyon¹ used drugs to prevent a migrating mesodermal cell from reaching its normal position. This cell, the muscle precursor cell M, normally turns on *mab-5* transcription after it has migrated into the posterior region of the embryo. When prevented from migrating, it expressed *mab-5* at the normal time anyway. In some experimental embryos, the M cell expressed a *mab-5* reporter gene while still in the anterior. A striking aspect of the *C. elegans* lineage is that structures can be built up from cells that originate in widely separated lineage branches. For example, some pairs of *mab-5*-expressing posterior cells arise from lineage branches that separate at the first embryonic cleavage, whereas other such pairs of cells are sisters. Thus, if the results found for the two cell types studied hold for other *mab-5*-expressing cells as well, a complex lineage-based regulatory mechanism must operate to bring about the relatively simple regional pattern of *mab-5* expression.

A second surprising finding comes from the report in *Development*, where Salser and Kenyon² show that, during postembryonic development, *mab-5* expression cycles on and off in two ectodermal cell

lineages (Fig. 2). Ectopic expression studies as well as examination of mutants were used to show that this complex expression pattern is necessary for specification of multiple cell fates.

Thus, in contrast to the view that *Hox* genes are turned on early in the regulatory hierarchy and remain on as markers to specify regional identity, *mab-5* is turned on and off multiple times throughout postembryonic development to influence the fates of multiple individual cells. In specification of the morphogenetic identities of peripheral sensory structures known as rays, mutations in *mab-5* and its neighbour in the *C. elegans* *Hox* cluster, the *Abdominal-B* homologue *egl-5*, do not result in any change in cell lineage or cell type. They only affect subtle aspects of cell differentiation that dictate the specificity of heterotypic cellular associations^{6,7}. Thus, here the genes appear to be modulated at the very end of the developmental pathway to guide the differentiation of individual cells, and may have terminal differentiation genes such as those encoding cell-recognition or cell-adhesion proteins among their targets.

How general might this dynamic picture of *Hox* action be? It is, in fact, remarkably consistent with results of a study of the role of the *Hox* gene *Ultrabithorax* (*Ubx*) in specifying two parasegments in *Drosophila*. Castelli-Gair and Akam⁸ have shown that, as for *mab-5*, details of the intricate spatial and temporal expression pattern of *UBX* protein in parasegments 5 and 6 underlie the unique patterns of cell fates within these two segments. Castelli-Gair and Akam argue that segment 'identity' in this instance is not specified by whether, to what level or in what combination of other proteins *UBX* is expressed within individual cells, but rather by a distinct segmental *Ubx* expression programme.

With these results^{1,2,8}, the apparent uniqueness of *Hox* genes disappears. Their regulation and action look no different from those of other classes of regulatory transcription factors. If these characteristics hold for other *Hox* genes and other organisms, they represent a considerable step forward in dispelling the mystery surrounding *Hox* genes by unravelling precisely how they are regulated and how they specify regional identity. □

Scott W. Emmons is in the Department of Molecular Genetics, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461, USA.

Travels on rugged landscapes

Mark Buchanan

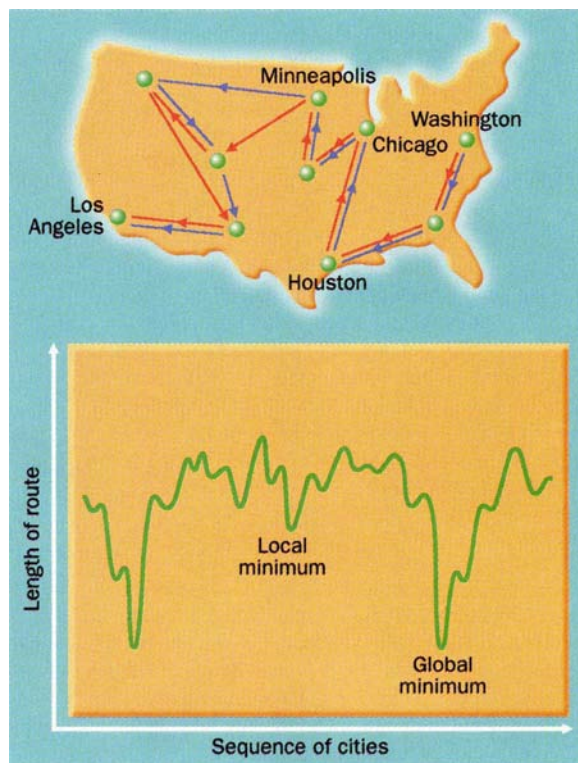
CREATIVITY in science, as elsewhere, begins with metaphor. In science, metaphor spawns theory, the ultimate value of which is judged by comparison with experiment. But metaphor also plays another role, as a vehicle for carrying powerful concepts and images from one area of science to

notions. At a conference held at Los Alamos in May*, physicists, chemists and biologists met to explore a metaphor that might serve to unify the study of these and other systems.

Mathematically, these systems have one common feature: their dynamics can be naturally represented by the motion of a point, representing the current state of the system, which climbs and falls over a complicated mathematical surface resembling a landscape. For some systems, the height of the surface represents the total energy: a rearrangement of particles — caused by thermal fluctuations, for example — might increase or decrease the system's energy, thus causing it to climb or fall on the energy landscape. In other problems, the landscape represents a different quantity, but plays a similar role. It might be a cost function, for example, which one would presumably want to minimize.

The most interesting dynamics occur on 'rugged' landscapes, which possess many local minima (see figure). A good example occurs in the theory of glassy materials. A system in thermodynamic equilibrium normally adopts a configuration which minimizes its free energy. Not so, however, for a glass. A glass-forming material, when cooled slowly from the liquid state, solidifies into an ordered crystalline structure. But if the liquid is instead cooled very rapidly, the resulting solid is characterized by locked-in microscopic disorder. This situation can be understood by supposing that the free-energy landscape is very rugged, and possesses

many local minima — configurations from which any small rearrangement would lead to an increase in free energy. These local minima correspond to the disordered glassy phase. Rapid cooling traps the system in this phase, and prevents it from reaching any of the global minima corresponding to the crystalline phase for



An old mathematical problem illustrates the notion of a rugged landscape. *a*, Suppose you are a travelling salesman, living in Washington DC, who must visit 10 specific cities in the United States, including Chicago, Houston and Minneapolis, and end up in Los Angeles. In which order should you visit the cities, so as to follow the shortest route? You might choose an arbitrary route like R1. To look for shorter routes, you might then consider all those which, like R2, differ from R1 by a single permutation of cities. One of these will probably be shorter than R1. Iterating the process, each time considering all one-permutation variants, you will eventually arrive at what seems to be a unique shortest route. But is it really the shortest? Part *b* shows the lengths of the various routes. Neighbouring routes differ by a single permutation. Starting from any point, the search strategy just described, by selecting shorter neighbouring sequences, marches inexorably towards the nearest minimum. If the landscape is 'rugged' — if the terrain fluctuates wildly and possesses many local minima — then such a local search strategy will lead inevitably to one of these, and will therefore fail to find the true shortest path.

another. Biological evolution, the mathematical theory of optimization and the physics of glassy materials might seem to have little in common. Nevertheless, they give rise to strikingly similar theoretical

1. Cowing, D. & Kenyon, C. *Nature* **382**, 353–356 (1996).
2. Salsler, S. J. & Kenyon, C. *Development* **122**, 1651–1661 (1996).
3. Lewis, E. B. *Nature* **276**, 565–570 (1978).
4. Lawrence, P. A. & Morata, G. *Cell* **78**, 181–189 (1994).
5. Peifer, M. et al. *Genes Dev.* **1**, 891–898 (1987).
6. Zhang, Y. & Emmons, S. W. *Nature* **377**, 55–59 (1995).
7. Chow, K. L., Hall, D. & Emmons, S. W. *Development* **121**, 3615–3626 (1995).
8. Castelli-Gair, J. & Akam, M. *Development* **121**, 2973–2982 (1995).

*Landscape Paradigms in Physics and Biology, 13–17 May, Los Alamos National Laboratory, New Mexico, USA.

which the free energy is truly minimized.

Under appropriate conditions, qualitatively similar and equally rugged landscapes characterize the penetration of magnetic flux into type-II superconductors¹, the packing of granular materials², the folding of proteins³ and the assembly of atomic clusters into amorphous structures⁴. Even the so-called travelling salesman problem (see figure) is not merely of academic interest, as it resembles problems encountered in the efficient routing of telephone connections, and in the placement of elements on an integrated circuit.

What makes these systems alike? Are there specific microscopic features, common to each, which underlie the existence of highly convoluted landscapes? This question seems to have been largely answered already, by developments in the theory of spin glasses⁵. A spin glass is a disordered material, the basic dynamics of which arise from the interactions of many magnetic moments. An example is a dilute solid solution of magnetic manganese atoms in copper. The random distribution of the manganese atoms causes the interactions between different pairs of atoms to be dissimilar. Some pairs have minimum interaction energy when parallel, others when antiparallel. As a result, the system struggles to find a unique configuration of lowest energy, and instead wanders between a large number of different but statistically similar 'near ground state' configurations. The inability of the system to find a unique ground state is referred to as 'frustration'.

So — frustrated dynamics lead to rugged landscapes. The next goal is to identify generic qualitative and quantitative features which might be exhibited by all frustrated systems.

On the theoretical front, one approach is to study simple mathematical models of rugged landscapes which are motivated by archetypal systems, and constructed on the basis of very general assumptions. In this way it is possible to derive, for 'typical' landscapes, the number of local minima and the expected time required for the system to evolve from its initial configuration to the nearest local minimum^{6,7}, where its evolution is arrested (A. Perelson and C. Macken, Los Alamos National Lab.). These results account well for the temporal evolution of antibodies during immune system response.

Studies of model atomic clusters⁴

(R. S. Berry, Univ. Chicago) suggest that energy landscapes can be usefully classified according to their topographies. Glass-forming materials appear to have many local minima (disordered metastable states) of roughly equal energy. On the other hand, in simple proteins and materials that form crystals, the landscapes are characterized by frequent large differences in energy between 'neighbouring' local minima. Whereas the former topography leads to the amorphous structures seen in glasses, the latter produces directed evolution towards the more 'focused' structures of crystals, and the specific compact folded states of proteins. In fact, the landscapes associated with proteins often appear to have funnel-like, self-similar structures (P. G. Wolynes, Univ. Illinois, and

J. N. Onuchic, Univ. California, San Diego) which help to guide the dynamical evolution rapidly towards the correct folded state.

It remains to be seen whether further work along these lines will lead to the development of a general theory of frustrated systems. But just as Feynman diagrams, developed originally in quantum electrodynamics, subsequently became essential tools in other fields such as statistical mechanics and fluid turbulence, so has the notion of the rugged landscape become a standard working image for physicists talking about atomic clusters, as well as for biologists speaking of protein folding. □

Mark Buchanan is an assistant editor at Nature.

EPIDEMIOLOGY

Links in childhood leukaemia

Sarah C. Darby and Eve Roman

LEUKAEMIA is the major malignancy of childhood in developed countries, accounting for around a third of all cancers diagnosed in children. Its incidence has a distinctive shape, with a marked peak at ages 2–3 followed by a steady decline (see figure overleaf). This pattern is consistent with the notion that exposures before birth and early in life may be important determinants of disease. For most cases the cause of the disease is uncertain. But recent scientific work has given rise to several hypotheses, and a paper by Petridou *et al.* (page 352 of this issue¹) points to another.

Genetic abnormalities, such as Down's syndrome, and also chemotherapy and radiotherapy, are well established causes of leukaemia. No more than 5% of leukaemias occurring in childhood can, however, be ascribed to such causes. In addition, there is considerable evidence from case-control studies, where the exposure of people with and without disease is compared, that diagnostic X-ray of the mother's abdomen during pregnancy can cause leukaemia; by contrast, cohort studies, where the disease rate in exposed and nonexposed individuals is compared, have failed to confirm this association, perhaps because of their low statistical power². Nowadays, it seems unlikely that *in utero* X-ray exposure could account for more than 1% of childhood leukaemias, because pregnant women are X-rayed less frequently and doses are much lower than in the past.

As well as the causes listed above, for which risk estimates can be quantified directly, *in utero* and postnatal exposures to an ever-growing list of biological, physical and chemical agents have been suggested

as risk factors for childhood leukaemia; among the suspects are electromagnetic radiation, hydrocarbons, pesticides, vitamin K and a variety of infections. Indeed, the epidemiological evidence that infectious agents are involved is considerable^{3,4}, but as yet no definite, or even candidate, agents have been identified.

It seems reasonable also to postulate that natural background radiation may cause a fraction of childhood leukaemias. The proportion that can be attributed to such radiation cannot, however, be estimated directly, because most children are exposed at very similar dose rates. Indirect estimates, based on extrapolation from data on survivors of the Japanese bombings, who were exposed at much higher doses and dose rates, indicate that around 7% of childhood leukaemias could be caused by postnatal exposure to background radiation, and risk estimates from studies of the effects of X-rays in pregnancy indicate that perhaps another 7% might be due to *in utero* exposure^{5–7}.

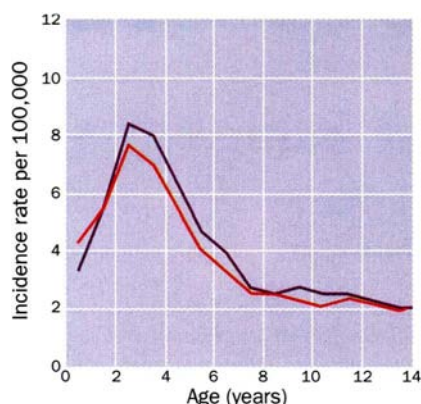
Occasionally, following events such as the accident at the Chernobyl nuclear plant in 1986 and the atmospheric nuclear weapons tests in the early 1960s, some populations receive much more low-dose-rate radiation than others. Until now, such studies have not reported evidence to contradict the estimates based on extrapolations from exposure to medical or atomic-weapon sources of radiation^{8,9}. But none has specifically examined infant leukaemia categorized by *in utero* exposure.

This is what Petridou *et al.* have done, focusing on the incidence of infant leukaemias in Greece following the

1. Fisher, M. P. A. *Phys. Rev. Lett.* **62**, 1415–1418 (1989).
2. Liu, C.-H. *et al. Science* **269**, 513–515 (1995).
3. Wolynes, P. G., Onuchic, J. N. & Thirumalai, D. *Science* **267**, 1619–1620 (1995).
4. Ball, K. D. *et al. Science* **271**, 963–966 (1996).
5. Binder, K. & Young, A. P. *Rev. Mod. Phys.* **58**, 801–976 (1986).
6. Perelson, A. S. & Macken, C. A. *Proc. Natl Acad. Sci. USA* **92**, 9657–9661 (1995).
7. Kauffman, S. A. *Origins of Order, Self Organization and Selection in Evolution* (Oxford Univ. Press, 1993).

accident at Chernobyl. Because of local weather conditions at the time, northern Greece experienced some of the highest levels of contamination from the accident outside the former Soviet Union. Estimates of the resulting average population dose during the following year are of the order of 1 millisievert (refs 10, 11), similar to that received annually from natural background radiation. If the risk estimates based on extrapolation are wrong, and a substantial proportion of infant leukaemias are caused by natural background radiation, then one would expect the Chernobyl exposure to cause an approximate doubling of risk in Greece.

That is indeed what is reported by



Incidence of childhood leukaemia in England and Wales, showing the peak at ages 2–3 and a steady decline thereafter (purple, male; red, female). In Europe, about 1 in 2,000 children develop leukaemia before their fifteenth birthday, but the cause of the vast majority of cases is unknown. (Source: Leukaemia Research Fund, University of Leeds.)

Petridou *et al.* — infants aged up to 1 year old who were exposed *in utero* to Chernobyl irradiation have a 2.6-fold increase in leukaemia compared with children born earlier or later. In addition, among Greek children exposed *in utero* to Chernobyl radiation, those born to mothers living in high radioactivity areas had higher rates than children born to mothers living in low radioactivity areas.

It is important, however, not to assume that the reported association is causal. First, there is no *a priori* reason to suppose that leukaemia in infants is more readily induced by *in utero* radiation than leukaemias in older children, and no increase was observed in older Greek children exposed *in utero* to Chernobyl radiation. Second, although there is a gradient in incidence of childhood leukaemia from low through medium to high radioactivity areas, the question arises as to whether those living in high radioactivity areas would actually have received the highest doses, since the majority of the Chernobyl exposure came from ingestion of contami-

nated foodstuff; moreover, no account is taken of the fact that exposures from Chernobyl lasted several years, so that children born after the period defined as 'exposed' would also have received some Chernobyl exposure.

Third, the results depend on only 12 cases of infant leukaemia occurring in children who were *in utero* during the high exposure period, of whom only four were born to mothers living in the high radioactivity areas at the time of diagnosis. Finally, cytogenetic techniques suggest that infant leukaemia constitutes a distinct subset of childhood leukaemias, with approximately 80% of cases having an abnormality at position 23 on the long arm of chromosome 11 (11q23)¹². But no diagnostic data are given on the type of leukaemia, or on how many children actually had the defect at 11q23.

It is to be hoped that Petridou *et al.* will be able to present at least some of this additional information elsewhere. Even if they do, however, as is usual in epidemiology, observational data from a single study cannot be regarded as any more than hypothesis-generating. Corroborating evidence is needed, for example from other groups of children exposed *in utero* to different levels of environmental radiation, before any firm conclusions can be drawn.

For all the drawbacks and difficulties of their study, Petridou and colleagues have provided a welcome contribution to the literature. The causes of childhood leukaemia are likely to prove many and complicated. The more data and hypotheses we have with which to attack the problem, the better. □

Sarah C. Darby is in the Imperial Cancer Research Fund Cancer Epidemiology Unit, Gibson Building, Radcliffe Infirmary, Oxford OX2 6HE, UK. Eve Roman is in the Leukaemia Research Fund Centre for Clinical Epidemiology, 17 Springfield Mount, Leeds LS2 9NG, UK.

- Petridou, E. *et al.* *Nature* **382**, 352–353 (1996).
- United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR) (UN, New York, 1988).
- Greaves, M. F. & Alexander, F. E. *Leukaemia* **7**, 349–360 (1993).
- Kinlen, L. J. *Br. J. Cancer* **71**, 1–5 (1996).
- Darby, S. C. in *Origins of Human Cancer: A Comprehensive Review* (eds Brugge, J. *et al.*) 183–190 (Cold Spring Harbor Laboratory Press, NY, 1991).
- US National Academy of Sciences Committee on Biological Effects of Ionizing Radiation *BEIR-V Rep.* (US NAS, Washington, DC, 1988).
- Bithell, J. F. & Stiller, C. A. *Stat. Med.* **7**, 857–864 (1988).
- Parkin, D. M. *et al.* *Br. J. Cancer* **73**, 1006–1012 (1996).
- Darby, S. C. *et al.* *Br. Med. J.* **304**, 1005–1009 (1992).
- United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR) (UN, New York, 1994).
- Papastefanou, C., Manolopoulou, M. & Charalambous, S. *J. Envir. Radioact.* **7**, 49–64 (1988).
- Pui, C.-H., Kane, J. & Crist, W. M. *Leukaemia* **9**, 762–769 (1994).

Ultra-infrared

THE great G. N. Lewis had the true experimenter's suspicion of laboratory instruments. A student of his once obtained an unlikely reading from their infrared spectrometer. Lewis took off the hood and looked into the beam. "Aha," he said, "I am the first man to see the infrared. And it is green!" Visible light was leaking into the optics.

Daedalus now plans to repeat this observation for real. He reckons that we could all see the infrared, were it not absorbed by the lens and fluid of the eye. Last week, he proposed to mix a clockwise, circularly polarized microwave beam at one frequency with a counter-clockwise beam of a lower frequency. Their summed field vectors should rotate clockwise at the difference-frequency. Suppose, he now says, you do this with light. Two superposed, visible, coherent, circularly polarized laser beams could easily have a difference-frequency in the infrared. But the retina is a nonlinear detector composed of biological, left-handed molecules. Hit by the mixed beam, it should detect the rotating difference frequency, and 'see' the effective infrared.

What colour will the infrared be? It might indeed be green — intense enough near-infrared can give a dim sensation of green. But it might be any other colour, or even some unimaginable new colour, depending on how it affects the retinal sensors. A whole range of such new colours could be generated by varying the frequencies of the original circularly polarized laser beams.

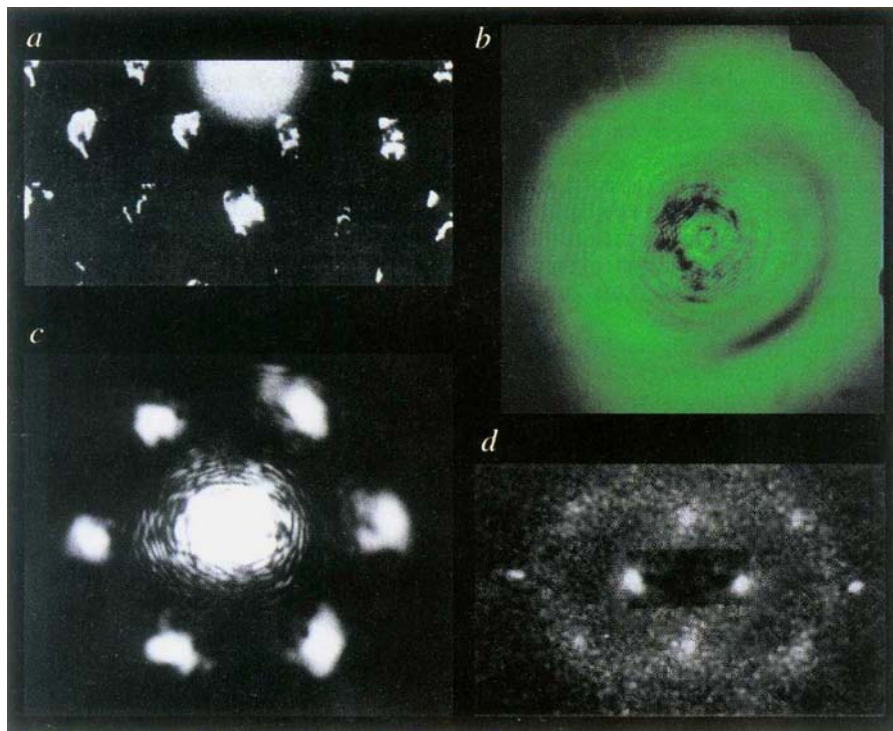
DREADCO's new circularly bi-polarized lamp, the 'Circlamp', will shed a searching new light on the materials around us. Living things and natural materials, made of left-handed molecules, show circular dichroism: they absorb clockwise and counter-clockwise polarizations differently. Under the Circlamp, they will shine out in a characteristic set of infrared colours. Synthetics, with their symmetrical or racemic molecules, will have quite a different chromatic range. The Circlamp will instantly spot synthetic fibres in an allegedly pure wool jacket, or nasty artificial chemicals added to organic foodstuffs, or synthetic dyes masquerading as natural colours.

Even physicists and electricians may benefit. A magnetic field can induce circular dichroism in many dyes. A sample of such a dye waved around under Circlamp light could map a magnetic field in contours of changing colour. Daedalus is now devising a special dichroic lacquer for motors, transformers, and so on. The Circlamp will then reveal their state of health. David Jones

Hexagonal patterns in laser light

SIR — In 1969, Schwarz and Hora reported an effect^{1–3} in which low-power laser light scattered through quartz produced hexagonal patterns (a in the figure). This effect has defied explanation and reproduction since its discovery and

We have carried out extensive studies⁵ in ferroelectric crystals with thermal focusing nonlinearities (in the simplest case, as with fluids, the density and hence index of refraction decreases in heated regions due to thermal expansion). We



a, Intermediate-field pattern of argon-ion laser light in quartz in the original Schwarz–Hora experiment¹; b, far-field pattern in PMN crystals (present work: 1/30-s exposure; dark ring is at 1.0°); c, potassium niobate⁵ pattern; d, colloidal suspensions¹⁴.

has often been regarded with scepticism, although recently Honda discovered striking hexagonal patterns in the laser light scattered from KNbO_3 (ref. 4). Here we report a similar, reproducible effect: low-intensity laser light scattered through crystals gives rise to hexagonal intensity patterns. We show that the phenomenon arises from thermal focusing in photorefractive media.

observe (b in the figure) far-field optical patterns in $\text{PbMg}_{1/3}\text{Nb}_{2/3}\text{O}_3$ (PMN), that contain interference rings with intensity distributed with a 6-fold symmetry around the beam direction. There is a threshold for this effect.

We find that these patterns do not rotate when the sample is rotated, ruling out crystallographic-diffraction explanations. Banerjee *et al.* report⁶ the complementary effect of rotation of the images with stationary geometry, which we believe to be due to photorefractive with transverse thermal instabilities⁷. Firth's relationship⁸ determines the wavevector K of the transverse instability: $K^2 L = 3\pi n k / 2$, where L is the path length; n , the index of refraction (2.4); and k , the wavevector of the laser ($19,436 \text{ cm}^{-1}$). Hence K is $2,850 \text{ cm}^{-1}$ and scattering angle $\Theta = \tan^{-1}(K/k) = 0.45^\circ$, about midway between the value 1.0° observed in KNbO_3 and 0.20° measured in PMN and comparable to that in the Schwarz–Hora effect.

Immediately after illumination, a circular ring is observed in the far-field

pattern of several ferroelectric oxides⁵, purely a thermal-lens effect arising from interference between light scattered from the centre and edges of the illuminated region. If the laser power is sufficient that the maximum thermal-lens diffraction angle exceeds that for pattern formation, Θ , then within a few seconds the ring coalesces into six spots. This shows that thermal defocusing provides the transverse nonlinearity necessary for pattern formation, although the observability of the pattern is intrinsically a photorefractive effect, caused by a change in refractive index due to the electric field of the light.

These phenomena can be described elegantly using the theory of two-dimensional melting. As noted by others^{9,10}, Kosterlitz–Thouless theory¹¹ predicts a two-step evolution of diffraction patterns: the second step is from a circular Debye–Scherrer ring to a hexatic phase with 6-fold symmetry and angularly broadened spots. In this interpretation the unilluminated specimens contain a random collection of charged defects. The standing light wave causes them to order ('freeze'), and spontaneous symmetry breaking in two dimensions causes the 'Debye–Scherrer' circular diffraction pattern to evolve into a hexagonal pattern. (We see this effect only in single-crystal PMN, not in ceramics where diffusion of defects through grain boundaries to produce the requisite holographic gratings is prevented.) At high laser intensity, pattern assembly becomes continuous. This satisfies the tricritical nature¹² of other nonequilibrium 'phase transitions'.

Schwarz and Hora mentioned¹ that the far-field patterns resembled the electron diffraction patterns in the target crystals; but our explanation offered here suggests that this resemblance was misleading¹³ — hexagonal patterns would still have resulted if cubic lattices¹⁴ had been used as targets.

J. F. Scott

*Faculty of Science,
University of New South Wales,
Sydney 2052,
Australia*

R. A. O'Sullivan

*Department of Applied Physics,
Royal Melbourne Institute of Technology,
Melbourne 3001,
Australia*

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. Priority will be given to letters of fewer than 500 words and five references.

- Schwarz, H. & Hora, H. *Appl. Phys. Lett.* **15**, 349–351 (1969).
- Harris, R. L. & Smith, R. F. *Nature* **225**, 502 (1970).
- Schwarz, H. *Nature* **225**, 1173 (1970).
- Honda T. *J. Mod. Opt.* **37**, 851–853 (1990).
- Chen, T., Scott, J. F. & Phillips, P. E. *Integ. Ferroelec.* **5**, 1–12 (1994).
- Banerjee, P. P. *et al.* *Opt. Lett.* **20**, 10–12 (1995).
- Grynberg, G. *Opt. Commun.* **66**, 321–324 (1988).
- Firth, W. J. *J. Mod. Opt.* **37**, 151–153 (1990).
- Nelson, D. R. & Halperin, B. I. *Phys. Rev. B* **19**, 2457–2484 (1979).
- Merrin, N. D. *Phys. Rev.* **176**, 250–254 (1968).
- Kosterlitz, J. M. & Thouless, D. J. *J. Phys. C* **6**, 1181–1203 (1973).
- Scott, J. F., Sargent, M. & Cantrell, C. D. *Opt. Commun.* **15**, 13–15 (1975).
- Macdonald R. & Eichler H. *J. Opt. Commun.* **89**, 289–295 (1992).
- Chowdhury, A., Ackerson, B. J. & Clark, N. A. *Phys. Rev. Lett.* **55**, 833–836 (1985).

Organization of the visual cortex

SIR — One long-standing question in developmental neuroscience is whether the functional architecture of the visual cortex is epigenetically induced by activity-dependent mechanisms or whether it is prespecified by genetic instructions^{1,2}. It is widely assumed that dependence on experience and variability of the individual layout of cortical columns are the hallmarks of activity-dependent self-organization. Bonhoeffer and co-workers have challenged the hypothesis that the pattern of cortical orientation columns arises by activity-dependent refinement of geniculocortical connections^{3,4}.

Orientation maps in area 18 of cat visual cortex that were forced to form independently for the left and right eye nevertheless exhibit an unexpectedly large degree of similarity. Godecke and Bonhoeffer argued that their results should be viewed as evidence for an innate predetermination of orientation preference, if orientation preference is determined by the pattern of geniculocortical connections at all⁴. Taking an opposing point of view, we point out here that the observed phenomenon can result from activity-dependent mechanisms if not only microscopic mechanisms, but also global constraints imposed by areal borders and retinotopic organization are taken into account.

It has long been suggested that the activity-dependent refinement of geniculocortical connections might follow dynamics analogous to symmetry breaking and pattern formation in physical

systems. The qualitative behaviour of pattern-forming systems generally depends on the ratio of system size and characteristic wavelength of the emerging structure⁵. An archetypical example is the Rayleigh-Bénard instability, where convection rolls form in a fluid layer to transport energy from a heated bottom to a cold top. A small container can accommodate only a few rolls (strong confinement), and repetitions of the experiment will lead to identical patterns. In a large container the emerging pattern of rolls will no longer be predictable⁶. The relevant length scale in area 18 is its mediolateral extension (about 3 mm; ref. 7). The characteristic wavelength of orientation columns is approximately 1.3 mm (ref. 8). Hence the system studied by Bonhoeffer and co-workers represents the strong confinement case. A repetition of the developmental dynamics should then display very similar stationary solutions as long as areal borders are constant. Nevertheless, orientation maps should vary between different animals because the compartmentalization of the cortical sheet in functional areas also exhibits considerable inter-individual variation⁷.

To investigate this suggestion we studied an idealized model for the development of afferent receptive fields under the influence of pattern vision⁹ in cortical areas of different shape. The emergence of orientation maps was simulated using sequences of random stimuli applied to a

single model retina. In this system, a reverse-suture experiment should be compared with a pair of simulations starting from identical initial conditions but using different random sequences of stimuli. In areas with large extension in both spatial dimensions, we found maps with only a general structural similarity (*a, b* in the figure). In narrow and elongated areas, however, simulations using different sequences of random stimuli led to virtually identical orientation maps (*c, d* in the figure). Because variants of this model system have successfully explained various aspects of cortical map formation, and because our simulations clearly show the impact of global geometric constraints on map formation, we conclude that precise restoration of orientation maps in cat area 18 is in fact expected, if orientation maps arise by activity-dependent refinement of geniculocortical connections.

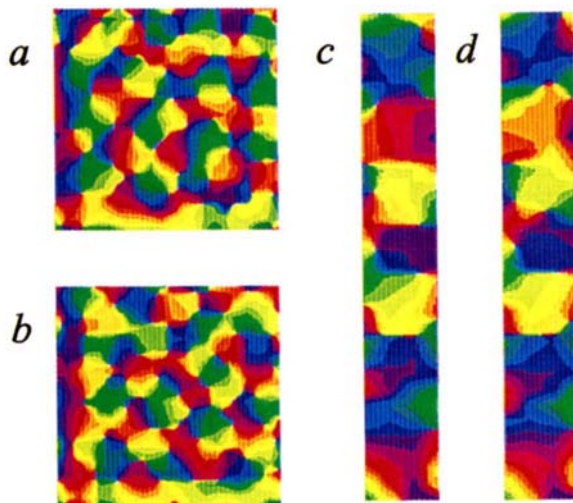
If external boundary conditions actually determine the development of area 18, the results of a reverse-suture experiment might be qualitatively different in a larger area where most columns are distant from areal borders. This is in fact suggested by two reverse-suture studies in area 17 (refs 10, 11) that are largely inconsistent with the results obtained in area 18. These studies indicate that neurons regaining input from a previously sutured eye exhibit independent orientation preferences in the two eyes.

It might therefore turn out that the phenomenon discovered by Bonhoeffer and co-workers does not represent evidence for an innate predetermination of orientation preference, but demonstrates the influence of global constraints on the formation of neural patterns.

F. Wolf, H.-U. Bauer, K. Pawelzik & T. Geisel

Max-Planck-Institut für
Strömungsforschung,
Postfach 2853, D-37018 Göttingen,
Germany
and SFB Nichtlineare Dynamik,
Universität Frankfurt,
Postfach 111932,
D-60054 Frankfurt, Germany.
e-mail: bauer@chaos.uni-frankfurt.de

Results of simulations of a simple model for the self-organization of orientation maps⁹ in a quadratic (*a, b*) and a strongly elongated target area (*c, d*). Pictures show spatial patterns of preferred orientations in a colour-coded representation. Orientation angles 0°, 45°, 90°, 135°, 180° are mapped on a colour circle: yellow-green-blue-red-yellow. Starting from identical initial conditions and using identical model parameters, the two self-organization processes differed only in the sequence of randomly chosen stimuli. In the first case the maps exhibit differing layouts



(mean correlation 0.28 ± 0.05 , evaluated for three pairs of maps); in the second case the maps are virtually identical (mean correlation 0.88 ± 0.04 , evaluated for three pairs of maps). Maps were generated using Kohonen's self-organizing map algorithm¹², with 16×16 (retinal) input channels mapped to 24×24 map neurons (*a, b*), and 16×8 input channels mapped to 36×6 map neurons (*c, d*), open boundary conditions. Stimuli were activity distributions of elliptic gaussian shape, length of half-axes $\sigma_1=1$, $\sigma_2=5$. Orientation and location of the stimuli were chosen by a numerical pseudo-random procedure, with different random sequences for *a* and *b*, and *c* and *d*, respectively. Further parameters of the algorithm were the adaptation step size $\epsilon = 5 \times 10^{-5}$, the width of the neighbourhood function $\sigma = 1$, and the number of adaptation steps $N = 3 \times 10^6$. Maps were initialized with pairwise identical, retinotopic connectivity patterns.

BONHOEFFER AND GÖDECKE REPLY — Wolf *et al.* make an interesting suggestion of how to explain our finding⁴ that in cat area 18 similar orientation maps emerge even if the eyes have no simultaneous visual experience.

Rather than contradicting our interpretation, we find that the notion that boundary conditions might determine the structure of orientation maps well in line with the idea that intrinsic or innate factors, rather than visual experience, control the exact layout of cortical orientation maps. In our paper we did not propose that activity-dependent mechanisms are

not needed; we argued rather that experience-dependent mechanisms might be less important than previously thought. In fact, we think that the determination of the structure of orientation maps by the geometry of a cortical area would be a good example of an intrinsic (rather than experience-dependent) determinant of the layout of orientation maps.

One might consider experimentally testing Wolf *et al.*'s prediction that the result of our experiment would be different in area 17, which has a different geometry. Unfortunately this idea is difficult to test, as in cats the main part of area 17 lies buried in the medial bank and is therefore inaccessible to optical imaging. The parts of area 17 lying on the surface of the brain are very close to the area 17/18 boundary. Because it has been shown that orientation maps are continuous at this border¹³ the optically accessible parts of area 17 will be influenced by area 18 and there-

fore orientation maps in these parts of area 17 are again likely to be stable in our experimental paradigm.

Tobias Bonhoeffer, Imke Gödecke
Max-Planck-Institute für Psychiatrie,
Am Klopferspitz 18a,
82152 München-Martinsried, Germany.

1. Wiesel, T. N. *Nature* **299**, 583–591 (1982).
2. Singer, W. *Science* **270**, 758–764 (1995).
3. Kim, D.-S. & Bonhoeffer, T. *Nature* **370** 370–372 (1994).
4. Gödecke, I. & Bonhoeffer, T. *Nature* **379**, 251–254 (1996).
5. Cross, M. & Hohenberg, P. *Rev. Mod. Phys.* **65**, 850–1112 (1993).
6. Gollub, J. & Steinman, J. *Phys. Rev. Lett.* **47**, 505–508 (1981).
7. Tusa, R., Rosenquist, A. & Palmer, L. *J. Comp. Neurol.* **185**, 657–678 (1979).
8. Lowel, S., Freeman, B. & Singer, W. *J. Comp. Neurol.* **255**, 401–415 (1987).
9. Obermayer, K., Ritter, H. & Schulten, K. *Proc. Natl Acad. Sci. USA* **87**, 8345–8349 (1990).
10. Blakemore, C. & Van Sluyters, R. *J. Physiol. (Lond.)* **237**, 195–216 (1974).
11. Movshon, T. *J. Physiol. (Lond.)* **261**, 125–174 (1976).
12. Kohonen, T. *Biol. Cybern.* **43**, 59–69 (1982).
13. Bonhoeffer, T., Kim, D.-S., Maloney, D., Shoham, D. & Grünvald, A. *Eur. J. Neurosci.* **7**, 1973–1988 (1995).

Feeding inhibition by neuropeptide Y

SIR — OB protein (leptin) is a circulating signal linking fat mass to the brain control of energy balance^{1–3}. Administration of OB protein to the brain induces long-lasting reductions in food intake and body weight of obese *ob/ob* mice^{1,2}. This striking finding, together with the observation that the circulating OB protein concentration is proportional to body fat in humans⁴, has stimulated research on the brain mechanisms involved in OB protein action. The inhibitory effects of OB protein on neuropeptide Y gene expression^{2,3} and secretion² have led to the proposal that neuropeptide Y, a potent stimulator of feeding⁵, is the major mediator of the actions of OB protein^{3,6,7}. To test this hypothesis, we examined the interaction of brain administration of OB protein and neuropeptide Y on the feeding behaviour

of *ob/ob* mice. We found that OB protein inhibits the expected feeding following administration of neuropeptide Y. This result indicates that OB protein can functionally antagonize the actions of exogenous neuropeptide Y and suggests that the receptor-mediated actions of this neuropeptide on feeding are under the control of OB protein.

We implanted chronic brain-infusion cannulas in the lateral ventricle in obese C57BL/6J *ob/ob* mice under anaesthesia. We gave free-feeding mice two intracerebroventricular injections; recombinant human OB protein (0.001–1 µg per mouse) or artificial cerebrospinal fluid contained in a volume of 1 µl (ref. 1) followed 1 h later by single doses of neuropeptide Y (0.1–5 µg per mouse) or cerebrospinal fluid.

The figure shows the cumulative food intake of these obese *ob/ob* mice. Food intake increases following neuropeptide Y injections (a), whereas it decreases following OB protein administration (b). However, when injections of OB protein precede those of neuropeptide Y, 4-h cumulative food intake significantly decreases less than or equal to that measured following artificial cerebrospinal fluid (c). Food intake following 5 µg neuropeptide Y is similarly inhibited by 0.1 and 1 µg OB protein (45 %), whereas 1 ng OB protein has no effect. Clearly, the magnitude and time course of food intake in these mice are determined by the presence of OB protein (c in the figure). Although the stimulatory actions of neuropeptide Y on feeding are dominated by OB protein, the increased food intake following combined OB protein and neuropeptide Y administration indicate that the neuropeptide is still acting at its post-synaptic receptors.

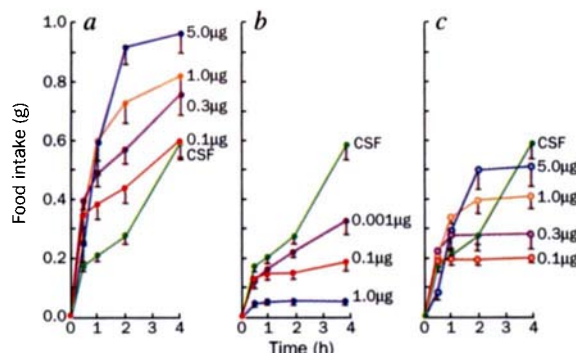
Our results suggest that OB protein determines the sensitivity of the feeding response to exogenous neuropeptide Y. OB protein, when placed into the lateral ventricle, may alter the binding of exogenous neuropeptide Y to its receptors that are critical for feeding and/or may inhibit downstream signal transduction^{5,8}. Thus, it appears that neuropeptide Y interacts, particularly in the early post-injection period, with OB protein or OB protein-dependent signalling processes contributing to the regulation of body energy balance^{9,10}. Nevertheless, our results are also consistent with other mediators of OB protein action, for example as suggested by a report that normal-weight mice lacking neuropeptide Y respond to peripheral administration of OB protein¹¹. Disruption of the usual balance and interplay between OB protein, neuropeptide Y and other unidentified mediators of OB protein action seem to be important factors in the aetiology of obesity and non-insulin-dependent diabetes mellitus^{5,8}.

Françoise J. Smith, L. Arthur Campfield, John A. Moschera, Pascal S. Bailon & Paul Burn

Preclinical Research and Development,
Hoffmann-La Roche Inc.,
Nutley, New Jersey 07110, USA

1. Campfield, L. A., Smith, F. J., Guisez, Y., Devos, R. & Burn, P. *Science* **269**, 546–549 (1995).
2. Stephens, T. W. *et al.* *Nature* **377**, 530–532 (1995).
3. Schwartz, M. W. *et al.* *Diabetes* **45**, 531–535 (1996).
4. Considine, R. V. *et al.* *New Engl. J. Med.* **334**, 292–295 (1996).
5. Frankish, H. M., Dryden, S., Hopkins, D., Wang, Q. & Williams, G. *Peptides* **16**, 757–771 (1995).
6. Ezze, C. J. *NIH Res.* **7**, 39–45 (1995).
7. Rohner-Jeanrenaud, F. & Jeanrenaud, B. *N. Engl. J. Med.* **334**, 324–325 (1996).
8. Kaiyala, K. J., Woods, S. C. & Schwartz, M. W. *Am. J. Clin. Nutr.* **62**, 1123s–1134s (1995).
9. Bray, G. A. & Campfield, L. A. *Metabolism* **24**, 99–117 (1975).
10. Leibel, R. L., Rosenbaum, M. & Hirsch, J. *New Engl. J. Med.* **323**, 621–628 (1995).
11. Erickson, J. C., Clegg, K. E. & Palmiter, R. D. *Nature* **381**, 415–418 (1996).

Effects of intracerebroventricular (ICV) administration of Ob protein on the feeding response to ICV doses of neuropeptide Y (NPY) in free-feeding *ob/ob* mice. Cumulative food intake during the 4 h after the second ICV injection (time=0) is shown. a, Increased food intake after doses of NPY alone ($n=8$); b, decreased food intake after doses of human Ob protein alone ($n=7-10$); c, interaction



of 1 µg per mouse human Ob protein and increasing doses of NPY ($n=5-9$). Cumulative food intake following injection of artificial cerebrospinal fluid (CSF; $n=25$) is shown in each panel for comparison. Food, but not water, was removed from the cage during the 1-h period between ICV injections. All experimental procedures were performed according to protocols approved by the institutional Roche animal care and use committee. Values, mean $\pm$ s.e.m. Statistics were tested using Student's *t*-test with unpaired samples with $P < 0.05$.

Function of 14-3-3 proteins

SIR — 14-3-3 proteins were first characterized as very abundant acidic proteins in the brain¹, functioning as kinase-dependent activators of tyrosine and tryptophan hydroxylase involved in the biosynthesis of neurotransmitters². More recent studies indicate that they are multifunctional regulators participating in cell-cycle control and various signal transduction pathways^{3,4}. Their functions are, however, still incompletely

In the process of characterizing the human 14-3-3 ϵ cDNA, we found unexpectedly that the 3' untranslated region of this cDNA showed a 285 bp match with clone 8-1 of the probe for Miller-Dieker lissencephaly syndrome^{6,7} (see figure). Because the 8-1 sequence aligned with the 3' untranslated region of the human 14-3-3 ϵ cDNA, we assessed the validity of this match by comparing the human sequence to corre-

the human 14-3-3 gene is probably contained in the 8-1 probe previously used to define the 5' end of the MDCR locus. Hence, the 8-1 probe is chimaeric for more than one gene, and the exact boundaries of the deletions associated with Miller-Dieker lissencephaly syndrome require further clarification.

Dong-Yan Jin

Myung Soo Lyu

Christine A. Kozak

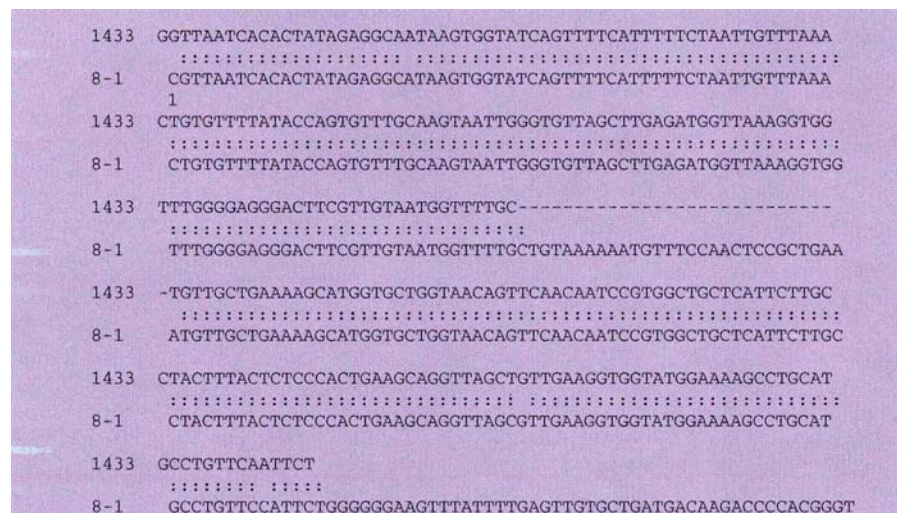
Kuan-Teh Jeang*

Laboratory of Molecular Microbiology,
National Institute of Allergy and Infectious
Diseases,

National Institutes of Health,

Bethesda, Maryland 20892-0460, USA

* To whom correspondence should be addressed. e-mail: kjeang@atlas.niaid.gov



Sequence alignment of the human 14-3-3 ϵ and MDS 8-1 probe. Dots indicate identity. Bars indicate gap. GenBank accession numbers of the aligned sequences: human 14-3-3 ϵ , U43399; MDS 8-1 probe, L13388. The presented alignment starts at nucleotide 1204 of human 14-3-3 ϵ .

understood.

We have been studying the molecular mechanisms by which 14-3-3 proteins interact with cellular and viral oncoproteins. In this process, we independently isolated a 1,706 base-pair (bp) human 14-3-3 ϵ complementary DNA containing an intact 3' untranslated region with a poly(A) tail (GenBank accession number U43399). The 14-3-3 ϵ protein sequence translated from this cDNA is identical to human 14-3-3 ϵ sequence recently deposited in the database⁵, for which the 3' untranslated region was not reported.

sponding regions in the mouse and rat 14-3-3 ϵ homologues. We found perfect alignment throughout the length of the three 14-3-3 sequences (including the 3' distal untranslated regions: human 14-3-3 ϵ , U43399; mouse 14-3-3 ϵ , Z19599; and rat 14-3-3 ϵ , D30739), strongly suggesting that the 3' untranslated region of the 14-3-3 ϵ gene is highly conserved.

This sequence information suggests that 14-3-3 gene maps close to or within the previously defined 8-1 Miller-Dieker lissencephaly probe. To confirm directly the map location of the 14-3-3 ϵ gene in the mammalian genome, we used as a probe a 3' cDNA fragment unique to the human 14-3-3 ϵ and divergent from other 14-3-3 subtypes to type the progeny of an interspecies *Mus spretus* genetic cross⁸. We typed progeny for a *Hind* III restriction-fragment length polymorphism of 14-3-3 ϵ which shows tight linkage to *Evi2*. We could not identify any recombinants for this gene in 88 progeny, indicating that, at the upper level of the 95% confidence level, these genes are separated by 0–3.3 centimorgan. This position is in good agreement with a previous study which positioned the mouse homologue of Miller-Dieker chromosomal region (MDCR) just proximal to *Evi2*⁹.

Based on our results, we conclude that

REINER AND CASKEY REPLY — We agree with the mapping data presented by Jeang *et al.* above, but disagree with the interpretation. Our results indicate that the complementary DNA 8-1 encoding LIS-1 (ref. 6) overlaps the 3' end of the 14-3-3 ϵ gene. Two methods support this finding: first, we isolated cDNA 8-1 from a human fetal brain library; and second, we have demonstrated the existence of this transcript in RNA from both human fetal brain and mouse fetal brain by reverse transcriptase-polymerase chain reaction (RT-PCR).

Because of the low expression level of the 14-3-3 transcript (tested by Northern blot analysis) and the existence of at least five LIS-1 transcripts, we used RT-PCR as a sensitive detection tool to investigate this gene product. As the presence of this rare transcript in more than one species may emphasize its importance, we amplified both human fetal brain RNA and mouse brain RNA. We found that the 14-3-3 LIS-1 overlapping transcript is detectable both in human and mouse (figure and experimental details available from O. R. on request).

Our initial genomic analysis of a human cosmid clone indicates that the 3' sequences derived from 14-3-3 are located in one exon. We are now examining these genes in the Fugu fish genome, which has the advantages of a high degree of conservation, extensive synteny and a compact size. Overlapping genes are a well-known phenomenon among viruses¹⁰, but have rarely been described in vertebrates and invertebrates (see refs 11–13 for some examples). As LIS-1 is a gene tightly controlled during development¹⁴, we are presently investigating the regulation of transcription of the 14-3-3-LIS-1 mRNA during development.

Orly Reiner

Department of Molecular Genetics,
The Weizmann Institute of Science,
76100 Rehovot, Israel

C. T. Caskey

Merck Research Laboratories
RY 80K Rahway, 0765 New Jersey, USA

- Aitken, A. *Trends Biochem. Sci.* **17**, 498–501 (1992).
- Ichimura, T. *et al. Proc. Natl Acad. Sci. USA* **85**, 7084–7088 (1988).
- Burbelo, P. D. & Hall, A. *Curr. Biol.* **5**, 95–96 (1995).
- Morrison, D. *Science* **266**, 56–57 (1994).
- Conklin, D. S., Galaktionov, K., Beach, D. *Proc. Natl Acad. Sci. USA* **92**, 7892–7896 (1995).
- Reiner, O. *et al. Nature* **364**, 717–721 (1993).
- Dobyns, W. B. & Truwit, C. L. *Neuropediatrics* **26**, 132–147 (1995).
- Kozak, C. *et al. Mammalian Genome* **6**, 142–144 (1995).
- Kurtz, A. & Zimmer, A. *Mammalian Genome* **6**, 145–146 (1995).
- Selivanov, N. A., Filipov, A. G., Efimov, V. P., Marusich, E. I. & Mesyanzhinov, V. V. *Biomed. Sci.* **1**, 55–62 (1990).
- Yanicostas, C. & Lepesant, J.-A. *Mol. gen. Genet.* **224**, 450–458 (1990).
- Shayiq, R. M. & Avdhani, N. G. *J. Biol. Chem.* **267**, 2421–2428 (1992).
- Tetsu, O., Kanno, R., Isono, K., Taniguchi, M. & Kanno, M. *Biochim. Biophys. Acta* **1305**, 109–112 (1996).
- Reiner, O., *et al. J. Neurosci.* **15**, 3730–3738 (1995).

Evolving evolvability

Stuart Kauffman

Climbing Mount Improbable. By Richard Dawkins. Viking: 1996. Pp. 300. £20.

THE curious thing about evolution, it has been said, is that everyone thinks he or she understands it. Richard Dawkins is no exception. *Climbing Mount Improbable*, his latest mission to convey to a wide audience his own understanding of the subject, is well wrought, graceful and a delight to read. Not the least of its pleasures is that it hints at a possible *rapprochement* between two traditionally opposing explanations for the evolution of living complexity: a core neo-Darwinian view based overwhelmingly on functional explanation in terms of natural selection, and an older but now minority structuralist view based on the intuition that many of life's building blocks have their own ordered properties that account for some of the order in organisms. Yet even the possible *rapprochement* broached by Dawkins fails to push the limits of how curious evolution really is.

The central image of the book is "Mount Improbable" — a multi-peaked 'fitness landscape' (which shows the degree of adaptation of an organism to its environment) with sheer cliffs on one side and smooth slopes and uplands towards the high country on the other. Evolution, Dawkins tells us (reasserting Darwin's gradualism), can slowly accumulate improbable forms by natural selection only along the easy passages up the smooth side of the mountain. If accident can produce simple harmonies, only natural selection can achieve intricate ones by clambering upwards (rarely downwards) until one of the peaks is attained.

Many chapters of the book take their lead from simple computer programs that generate structures and forms through simple rules encoded by an underlying 'genome'. Topics covered include spider's webs and the associated design problems that must be solved. The chapter entitled "Museum of all Shells" builds on Daniel Dennett's image of a high-dimensional space of all organisms, each a morphological neighbour to those near it, and David Raup's theory of shell forms. The ontogeny of individual organisms and the evolution of form are trajectories in such a space. Simple programs that tune spiralling in the plane, widen the rate of the growing cone and extend growth into the third dimension give rise to a family of shells, many but not all of which correspond to existing ones. "Getting off the Ground" discusses various

pathways to gliding or flight itself. "The Forty-fold Path to Enlightenment" details possible alternative gradual pathways to the repeated independent evolution of sight using common variations in molecular or morphological features.

In a chapter entitled "Kaleidoscopic Embryos", Dawkins pursues the evolution of evolvability itself. He makes the plausible suggestion that symmetries and segmentation in embryos allow a mutation to influence all or specific subsets of the symmetries or segments. In this way,



Field of dreams — sunflowers are just one of many plants modelled by developmental computer programs in *The Algorithmic Beauty of Plants* by Przemyslaw Prusinkiewicz and Aristid Lindenmayer. Springer, DM48, \$29.95 (pbk).

a single mutation in the fruitfly *Drosophila* may have bilateral symmetric effects in all segments at once, or in the thoracic subset of segments alone, or in a single segment. These symmetries make it easier for natural selection to 'find' useful mutants affecting many parts of the embryo simultaneously, and so ease adaptation.

Dawkins discusses briefly the structuralist tradition of something like 'natural forms'. Although admitting that the extent to which such forms inform evolution is an open matter, he insists that natural selection alone can account for highly improbable assemblages of simple building blocks. This stance is reasonable, but understates the main issue: why

should Mount Improbable happen to have any smooth sides in the first place? More broadly, what properties characterize complex systems so they can be assembled by an evolutionary, indeed a coevolutionary, process? It is fair to say that we hardly know.

Are the smooth sides up which tinkering natural selection wends its way the natural properties of the molecular and morphogenetic building blocks of organisms? Or, because selection cannot otherwise succeed, are they themselves an achievement of evolution? Most probably, both are the case. So the very *a priori* plausibility of Dawkins' discussion of the evolution of flight and sight suggests that the appropriate properties 'lie to hand', given the earlier evolution of developmental mechanisms yielding reliable form. But what can those mechanisms

be? Proponents of a drastic genetic determinism, the common view, roughly picture the genome as a computer program able to generate virtually any output. Yet DNA can specify only RNA and, with the help of a regulatory network, the time and place of RNA synthesis. The rest of cell behaviour and morphogenesis, from the self-organization of cell membranes to the unfolding of limb formation, is, bluntly, up to physics and chemistry. If these assembly processes were extremely improbable and hard to maintain, they would not have been discovered or sustained by evolution.

Developmental mechanisms that yield robust morphogenesis simultaneously yield smooth fitness landscapes. One

probable example is spiral phyllotaxis (that is, the spiral arrangement of lateral organs such as leaves, scales or florets) in pine cones, sunflowers and many other plant structures. The apical meristem gives rise to spirals turning clockwise and spirals turning anti-clockwise, with the numbers in the two sets being adjacent integers of the Fibonacci series: 3:5, 5:8, 8:13. These patterns can be expected as the stable bifurcation patterns in a model system. Each broken symmetry tends to 'choose' the way in which the next symmetry will be broken, so the pattern is robust. Although the detailed model may not be right, it exemplifies the concept of natural forms, readily and robustly accessible from a class of developmental mechanisms.

A second probable example is the emergence of orderly behaviour in models of parallel-processing genetic regulatory networks. After three decades of work, biologists, mathematicians and physicists are converging on the conclusion that such networks typically behave in two broad regimes, ordered and chaotic, with a phase transition in between. The ordered regime is readily accessible by tuning simple general construction features of genetic networks, such as the numbers of regulatory inputs per gene, and simple biases on the class of control rules governing the activities of regulated genes as a function of their inputs. Study of genomic systems to test for these biases is in progress. Meanwhile, the typical properties of networks in the ordered regime seem to parallel many features of real ontogeny while remaining robust to mutational alteration in network structure and logic. Such systems therefore evolve on smooth landscapes.

The evolution of evolvability is a deep and subtle problem. Not all complex systems can be assembled by an evolutionary process; their assembly corresponds to the cliffs of Mount Improbable. Further, recent theory indicates that for evolution to succeed in assembling complex systems, the structure of fitness landscapes must be correlated with the 'search procedures' guiding evolution on these landscapes. Averaged over all possible fitness landscapes, no search procedure seems to outperform any other. In particular, it seems that recombination is useful, and will be selected for, only on smooth multi-peaked fitness landscapes. Many species are sexual, presumably because recombination is a useful search technique. Again, where do such smooth landscapes come from? These observations hint at a collaboration between natural selection and the robust, readily available properties of organisms: niches easily explored and 'mastered' by the search procedures of mutation and recombination emerge jointly and self-

consistently with the organisms occupying these niches. We hardly understand this joint coevolutionary assembly of organisms, search procedures and readily explorable niches. Evolution is indeed curious. □

Stuart Kauffman is at the Santa Fe Institute, 1660 Old Pecos Trail, Santa Fe, New Mexico 87501, USA.

Life cycles

Nils Chr. Stenseth

Do Lemmings Commit Suicide? By Dennis Chitty. *Oxford University Press: 1996. Pp 268. £37.95, \$46.95 (hbk); £18.95, \$24.95 (pbk).*

DENNIS Chitty is known primarily for the 'Chitty hypothesis': "all species are capable of limiting their own population density without either destroying the food resources to which they are adapted, or depending upon enemies or climatic accidents to prevent them from doing so". A corollary of this broad hypothesis is that the periodic fluctuations in population size of several northern vertebrates such

tion) put forward to explain the so-called microtine cycle are expressed in fuzzier language and so are harder to confirm or refute. However, as an explanation of the lemming and vole cycle, the Chitty hypothesis has, as Thomas Huxley might have said, been "slain by the ugly facts". The kind of genetic polymorphism suggested by Chitty in the late 1950s and early 1960s, though, might be important in shaping the detailed anatomy of the lemming cycle.

Chitty had more success in his advocacy of the experimental approach for testing hypotheses about population fluctuations. This 'experimental reasoning' has had considerable influence on the ecological work carried out at the Bureau of Animal Population in Oxford. Although Chitty did not report on many controlled experiments, his students did. One of these was Charles J. Krebs, who crystallized the experimental method within the field of animal ecology, particularly in the study of fluctuating northern animal populations. The Chitty hypothesis has been tested extensively and as a result generally rejected. In this way Chitty has for more than half a century stimulated much of the discussion and scientific work on the microtine puzzle. The relatively few papers that he has written have been widely read; two have become citation classics.

Mary Evans Chitty's lifetime contribution to science is the subject of this autobiography, which tells the story of a scientist and teacher who clearly enjoys his work despite the scepticism with which it has been met. To quote David Hull in the foreword, Chitty "shows that a life in science without a Nobel Prize is still worth living". The book takes us from 1935, when Chitty arrived in Oxford to work with Charles Elton for a year (and ended up staying there for 26), right up to the present. All ecologists working on rodent cycles would be advised to read this book, as it gives insight into the sociology of this particular branch of ecology. It is the perfect sequel to Peter Crowcroft's *Elton's Ecologists: A History of the Bureau of Animal Population* (University of Chicago Press, 1991), and, as an account (somewhat biased) of one person's involvement, it represents a valuable historical document.

Curiously though, many historically important contributors to the field (such as the Finnish biologist Olavi Kalela) are either not mentioned at all or treated rather superficially. Indeed, Chitty comes down rather hard on many colleagues who have criticized him or favoured different approaches and hypotheses. He admits, however, that he has behaved in much the same narrow-minded way, but

IMAGE
UNAVAILABLE
FOR COPYRIGHT

Lemmings do migrate, but not to the sea to commit suicide.

as the Scandinavian lemming are generated and sustained by a genetically determined behavioural polymorphism, with some genotypes favoured during peak densities, others during declining densities and the low-density phase.

This is a beautiful hypothesis based on fairly clear assumptions that can be tested experimentally. Most other hypotheses (involving, for instance, food and preda-

was unable to see it at the time. As he says, "most of one's time is spent in confusion, much in error, and a little with truth". This honest book is Chitty's sacrificial lamb.

The volume is dedicated to the late Helen Chitty, and is a valuable reminder of her significant contribution to the study of population cycles. Among other duties, she organized much of Elton's data by doing what Dennis Chitty calls "mail order zoology", including the Canadian Arctic Wildlife Enquiry and the Snowshoe Rabbit Enquiry, two important sources of information on spatio-temporal ecological variation.

There are, unfortunately, too many ugly, out-of-date 'facts' in the book that are not really synthesized into an integrated whole. The book could have done with further critical scrutiny and the use of an editor's red pen. Another problem is that it was written for three very different kinds of readers: those seeking what Chitty refers to as the ecological Holy Grail ("to understand the 10-year cycles in numbers of animals such as lemmings, voles, snowshoe hares, game birds and

defoliating insects"); scientists, whatever their speciality, who may be interested in the principles of scientific inference; and "anyone interested in mysteries, whether scientific or otherwise". Chitty is certainly aware of the difficulty of combining these three kinds of books into one: "in trying to kill two birds with one shot I here run the risk of merely wounding both". Perhaps he should have focused on just one type of reader: the scientist with an interest in the history of ecology and the development and discussion of his hypothesis.

The Chitty hypothesis may not be as beautiful as many long believed, and Chitty still does believe, but he has produced a very beautiful book. It is well written and full of interesting historical information. To give it colour there are quotations and references from a variety of sources, including the Bible and philosophy texts; indeed, seeing who he quotes is a fascinating exercise in itself. □

Nils Chr. Stenseth is in the Department of Biology, University of Oslo, PO Box 1050 Blindern, N-0316 Oslo, Norway.

Full circle

H. K. Moffatt

Turbulence: The Legacy of A. N. Kolmogorov. By Uriel Frisch. Cambridge University Press: 1995. Pp. 296. £45, \$80 (hbk); £15.95, \$29.95 (pbk).

THE phenomenon of turbulence holds a peculiar fascination both for physicists, who, like Einstein, still see it as the most challenging unsolved problem of classical physics, and for mathematicians, who see it as presenting the ultimate problem in the theory of dynamical systems. But for the meteorologist, oceanographer or engineer, turbulence holds a rather different kind of fascination as usually the unavoidable and intractable phenomenon at the heart of a given physical situation. It must be modelled through judicious approximations, based at best on physical insight or at worst on the need to come up with numbers, or through predictions, which may be no better than 'best guesses' in the absence of a satisfactory underlying theory for turbulent flow. In practical engineering, any small improvement in the modelling of turbulence could have important economic consequences, and this lends an edge to the urgency and competitiveness of research on turbulence. Nevertheless, progress at a fundamental level has been slow over the past 40 years, a reflection of the intense difficulty of the subject.

Uriel Frisch takes as his focus the theory of turbulence put forward by A. N. Kolmogorov in a famous series of three short papers that appeared in 1941 in *Comptes Rendu de l'Académie des Sciences de l'U.R.S.S.* These papers, published in Russian, were translated into English for the Aeronautical Research Council, and were 'discovered' by Batchelor shortly after the Second World War. He reviewed them, together with the almost parallel theories of Onsager, von Weizsacker and Heisenberg, at the Sixth International Congress of Theoretical and Applied Mechanics, held in Paris in 1946 (see *Nature* **158**, 883).

Kolmogorov's theory is based on two general hypotheses concerning the statistics of the relative velocities at a set of points in a turbulent flow. The net effect of these hypotheses is to identify a single dimensional parameter, namely the mean rate of dissipation of energy per unit volume, denoted ε by Kolmogorov (a notation that was subsequently adopted universally), as the key parameter that, together with the kinematic viscosity of the fluid ν , determines all the statistical properties of the small-scale ingredients of the turbulent flow. This theory flourished during the 1950s and early 1960s, when observational evidence from studies of turbulence in the atmosphere and ocean was found to agree substantially with the theory, at least as far

An arctic loon or diver, *Gavia arctica*, surveys a northern lake. The picture is one of more than 120 that appear in *Loons* by Aubrey Lang and Wayne Lynch. In the accompanying text the authors examine the scientific facts behind the many myths that have long surrounded these birds. Firefly, \$19.95 (pbk).



as the energy spectrum of turbulence governed by Kolmogorov's $k^{-5/3}$ law was concerned.

But Kolmogorov's theory contained the seeds of its own destruction, as was apparently recognized by Landau in discussions following the presentation of Kolmogorov's papers to the Russian Academy. The local rate of dissipation of energy per unit volume $\bar{\epsilon}$ is itself subject to spatial fluctuations, and, in subregions of the flow where $\bar{\epsilon}$ is much greater than its spatial average, the local statistics should reflect this local value rather than the mean value $\bar{\epsilon}$. Indeed, in regions where $\bar{\epsilon}$ is large, the process of cascade of energy from large to small scales is more intense, smaller scales of motion are generated, and the rate of viscous dissipation becomes even larger. This leads to the problem of intermittency of turbulent dissipation, a problem that was recognized by Kolmogorov himself in 1962 (*Journal of Fluid Mechanics* 13, 82–85) and has since attracted increasing attention.

All of this is discussed by Frisch in a stimulating and imaginative manner. He begins with a qualitative discussion of symmetry-breaking instabilities of a typical flow as the Reynolds number (Re) is increased from small values, and argues that in the asymptotic turbulent regime of very large Re , although all symmetries are broken in detail, a new statistical symmetry (essentially homogeneity and isotropy) emerges in relation to the small scales of motion. He then presents a novel discussion of the vital process of energy transfer (or 'cascade') from large to small scales. This leads to a discussion of the seeds of chaos in simple dynamical systems, and to a sophisticated introduction to some of the probabilistic techniques used in the theory of turbulence.

In the heart of the book, Frisch presents and develops his own reinterpretation of Kolmogorov's theory and contemporary theories incorporating the effects of intermittency. He bases his presentation on two empirical laws, the famous 'two-thirds law' relating to mean-square of the velocity difference between two points in the flow, and the 'law of finite energy dissipation', which states that $\bar{\epsilon}$ tends to a constant independent of ν as ν tends to 0. The first law is well supported by experiments, the second less so; if this second law is not valid — a possibility that Frisch does not care to contemplate — the whole Kolmogorov construction is a house built on sand.

One of the most illuminating discussions of intermittency is that involving the ' β -model' introduced by Frisch, Sulem and Nelkin in 1978. As the energy cascades, the volume in which 'active eddies' are present is supposed to reduce each step of the cascade by a factor $\beta < 1$; this leads to a fractal distribution for $\epsilon(x)$ from which higher-order statistics can also be deduced. Unfortunately, the theory is not in accord with subsequent experiments; to achieve agree-

ment, as Frisch shows, it is necessary to adopt a 'multifractal' rather than a simple fractal viewpoint, involving a continuum of coexisting fractal dimensions. At this point, many more earth bound researchers will part company with Frisch (and other multifractalists), unprepared to accept that the ultimate nature of turbulence may hinge on such an alien concept.

Much more promising, I believe, is the more conventional and physically appealing picture of turbulence as a random superposition of stretched vortex filaments, the 'sinews of turbulence'. Direct numerical simulations have been pointing in this direction for some years, and experiments by Couder and others (Douady, Couder and Brachet, *Physical Review Letters* 67, 983; 1991) lend credence to this view. The wheel of turbu-

lence is coming full circle as we witness a return, albeit at a highly sharpened level, to a consideration of concepts (vortices and vortex interaction) with which Kelvin would have felt at ease.

Frisch concludes his monograph with a long chapter providing a guided tour of further reading in the subject, an extremely useful aid for students whose appetites may have been whetted by this personal, highly unconventional and provocative account. This book will be a source of inspiration for theoretically inclined researchers, but will provide little comfort for practising engineers. □

H. K. Moffatt is in the Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Silver Street, Cambridge CB3 9EW, UK.

Holes unearthed

Stephen Battersby

Gravity's Fatal Attraction: Black Holes in the Universe. By Mitchell Begelman and Martin Rees. *W. H. Freeman: 1996.* Pp. 246. £19.95, \$32.95, £19.95.

"It's black, and it looks like a hole. I'd say it's a black hole."

SIDNEY Harris's white-coated cartoon astronomer makes this deadpan satire on scientific coinage and deduction while showing her colleague a picture of a star field with a big round gap in it. Although we do not yet have any images that are so

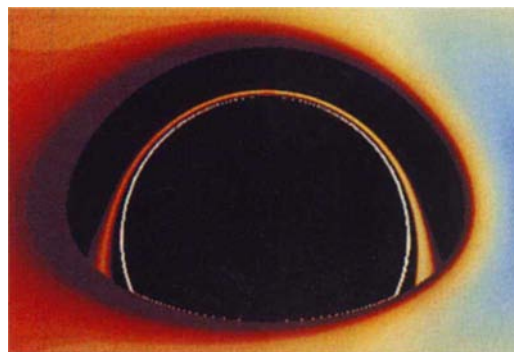
Exemplifying the book's confident stance, they say: "But we now know that black holes are not mere theoretical constructs; they exist in profusion and account for many of the most spectacular astronomical discoveries of recent times".

This certainty is prompted by a discovery that was made shortly before the book went to press. In the nucleus of the galaxy NGC4258, there are gas clouds in which water vapour forms natural masers. They trace out a warped disk of material orbiting at enormous velocities, betraying such a dense central concentration of mass that "the only plausible alternative is a single massive black hole". Unfortunately, the figures that illustrate this are by far the most confusing in the book; perhaps the news was so hot that they had to be printed at high speed. A few minutes of sweaty decipherment smudged the ink on two of them, strengthening my suspicion.

Other than this, there is little to criticize here and plenty to delight in, from the history (apparently one Revd John Michell was the first to postulate black holes in 1784, beating Laplace by ten years) to the well illustrated oddities of general relativity. Rigorous and thorough but not dry, the book describes holes of all shapes and sizes — all sizes, anyway — from the still-improbable primordial subatomic holes, shining by Hawking radiation, to the billion-solar-mass holes that power quasars.

Black holes have lost part of their traditional mystique — the question of their existence — but I am sure there is plenty left to keep them popular. □

Stephen Battersby is an assistant editor at Nature.



Simulated close-up view of a black hole and its disk of accreting gas, relativistically distorted by gravity and the high orbital speeds.

unequivocal, most astronomers assume that black holes probably do exist, either on the weight of circumstantial evidence or because their colleagues tell them so. But since 1963, when Roy Kerr calculated the properties of spinning black holes, we have been in the odd position of knowing exactly what a whole class of objects is like without knowing whether a single one exists.

When will the evidence accumulate to a point beyond reasonable doubt? It already has, according to Begelman and Rees.

Synthesis of inorganic materials with complex form

Stephen Mann & Geoffrey A. Ozin

Recent developments in inorganic materials chemistry suggest that concepts such as morphogenesis, replication, self-organization and metamorphosis could be useful for devising new synthetic strategies. Inorganic materials with complex form can be chemically synthesized by pattern replication of self-organized organic assemblies, such as micelles, vesicles and foams.

THE patterning of synthetic inorganic materials, exemplified by the complex, ordered pore networks of zeolites and other crystalline microporous materials, has traditionally been restricted to the nanometre scale (less than 1.5 nm). In these solids, the patterning of the pore architecture is achieved through the spatially controlled assembly of inorganic building blocks, facilitated by cationic or neutral organic molecules. In general, the templating function of the organic additive is to fill space, direct structure and balance charge¹. Recently, the imposition of patterns (pore networks) at a scale larger than the molecular has been achieved by using surfactant aggregates as templates that undergo cooperative self-assembly with inorganic species to form mesoporous materials with channel diameters of 1.5–10 nm (refs 2, 3). Structural patterning in inorganic materials synthesis at this scale still falls a long way short of the organic template-directed processes that result in the complexity of form and structure in biological minerals such as bone, shells and teeth.

Many biominerals are organized from the nanoscale to the macroscopic scale (a loose definition used here is: nanoscopic, <1.5 nm; mesoscopic, 1.5–100 nm; microscopic, 0.1–100 μ m; and macroscopic >100 μ m) to give hierarchical materials that are sculpted with complex form—spirals, spheroids and skeletons, for example—with apparent disregard for the rigid geometric symmetry of their inorganic constituents. Mackay⁴ has recently described the ‘flexicrystallography’, structure and symmetry properties of about 50 periodic minimal-energy surfaces, and how they allow one to transcend the boundary of traditional crystallography and the 230 classical space groups, to encompass periodic curved surfaces found in soft structures such as liquid-crystal mesophases and biological assemblies. This concept may provide a unifying connection between pattern formation in biomineralization and the sculpting of inorganic materials with complex form by synthetic organic templates.

The motivation for chemically constructing materials with meso- and microscopic form centres around the importance of shape and texture in determining properties such as flow and transport behaviour, catalytic activity, separation efficiency, adhesion, and storage and release kinetics in ‘smart’ colloids. Synthesizing inorganic materials with complex patterns could therefore be relevant to the design of new types of catalyst supports, membranes for the separation of large polymers, colloids and cells, biomedical implants with macroporosity, drug carriers, and vectors for delivery and release of viruses and DNA in transfection procedures. Acoustic behaviour, and heat- and mass-transport properties could be influenced by surface patterning or the fabrication of hollow and cellular structures.

To realize the perceived benefits of complex form in synthetic inorganic materials, it will be necessary to control the chemistry responsible for the ‘synthesis with construction’ of inorganic and

organic components on a series of length scales⁵. In this regard, biomineralization has much to offer in providing ideas and inspirations that can be developed in the synthesis of biomimetic materials^{6–11}. Here we focus on synthetic approaches related to biomineralization, illustrating how preorganized or self-organizing vesicles can expand the scale of inorganic materials patterning. An important aspect of biomineralization is that self-assembled organic templates are transformed by materials replication into organized inorganic structures. We show how this strategy can act as a guideline for synthesizing inorganic materials with complex form. The use of nonlinear diffusion-based processes¹²—typical of snowflakes and electrochemical deposits—or processing methods involving microlithography or physical moulding^{13,14} is excluded from our discussion because our objective is to identify potential chemical strategies for the ‘one-pot’ synthesis of patterned inorganic materials. Our aim is not to provide a comprehensive review of a new field of materials chemistry, but to stimulate further development by offering an agenda for the future.

Pattern formation in biomineralization

Although the complex form of biological minerals was of great interest to the nineteenth-century morphologists¹⁵, and later to D’Arcy Thompson¹⁶, the genetic basis for the diversity and evolution of biomineral patterns remains unknown. From a materials

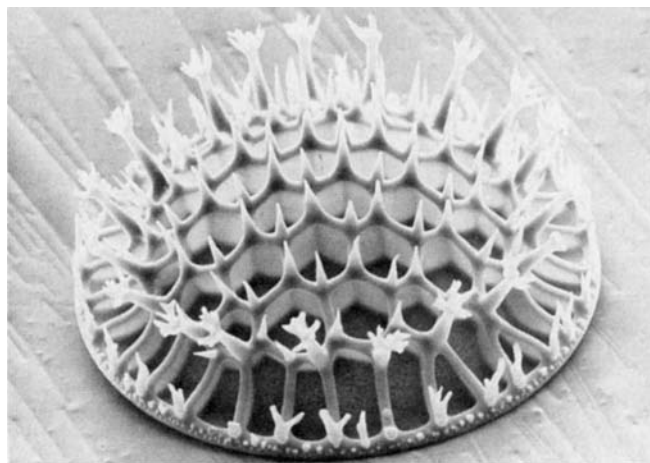


FIG. 1 Single valve of an unidentified diatom found in a deep-sea core of an Eocene marine deposit⁴⁹. (Picture courtesy of Phil. Trans.)

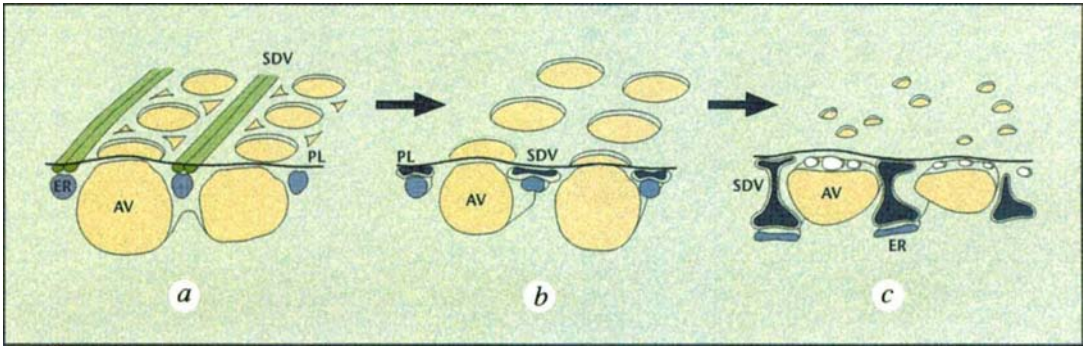


FIG. 2 Illustration of the important stages in the formation of the siliceous diatom exoskeleton. a, Silica deposition vesicles (SDV) are preorganized with microtubules around the boundary spaces of large areolar vesicles (AV) attached to the plasmalemma (PL). b, The SDVs are mineralized with amorphous silica (dark blue) to give a patterned porous wall. c, The

mineralized wall is thickened by extension of each SDV in association with the endoplasmic reticulum (ER). Detachment and retraction of the areolar vesicles from the plasmalemma results in infiltration with new SDVs and further mineralization of the pore spaces. Adapted from ref. 18.

perspective, the delicate filigree microskeletons of diatoms and radiolarians, or the spirals and cones of foraminiferal shells (Fig. 1), conflict with a commonplace view of inorganic minerals as rigid, inert, immutable materials of limited form and fabric. The complexity of these biomineral forms cannot be accounted for simply on the basis of inorganic processes of crystallization. Instead, they are patterned by assemblages of vesicles and other bilayer structures that are arranged and rearranged by cellular processes.

For example, the perforated silica shells of diatoms and many radiolarians are derived from close-packed arrays of large (areolar) vesicles that are secreted and attached to the membrane wall (plasmalemma) of the cell before mineralization (Fig. 2)^{17,18}. The areolar vesicles are not mineralized but their arrangement into a thin polygonal foam provides a patterned template extending to the meso- and microscale. In the diatom *Coscinodiscus*, a thin tubular membrane system is secreted from the Golgi apparatus and assembled along with microtubules in the gaps between the areolar vesicles¹⁸. Silica deposition is then confined tangentially within these tubular vesicles such that an open geometric mesh of mesopores is replicated (Fig. 2). Although this architecture often represents the final microskeletal form in radiolarians, diatoms often continue to process the areolar spaces with a delicate nanoporous pattern of silica. This is achieved by detachment and withdrawal of the areolar vesicles from the plasmalemma, and infiltration of the new interface with cytoplasm containing silica deposition vesicles and small unmineralized Golgi vesicles (Fig. 2). Together these vesicles self-assemble to produce a species-specific pattern of nanopores across the surface of the diatom shell.

An important determinant of the complexity of pattern and form expressed by biominerals is scale¹⁶. Different length scales bring into operation different biological processes. For relatively small-scale structures, such as the curved silica rods of choanoflagellates, vesicles can be shaped before mineralization by attach-

ment to the cytoplasmic membrane using two stress filaments¹⁹. In contrast, large architectures (coccoliths²⁰, for example) are usually constructed by patterning of the vesicles during biomineralization. Changes in shape arise as the mineral develops, and it is possible that the evolving pattern of morphogenesis results from a synergism between crystallographic and biological forces⁴. For multicellular organisms, the positioning and movement of cohorts of cells can generate patterns on a micro- and macroscopic length scale¹⁶. For example, the triradial symmetry of spicules formed in calcareous sponges is initiated from a cluster of three sclerocytes²¹. Continuous assemblages in the form of foams and rafts of closely packed cells facilitate the assembly of the labyrinthine skeleton of the crystalline shell (test) of echinoderms²². The spatial positioning of epithelial cells in shell mantle²³, osteoblasts and osteoclasts in bone matrix²⁴, and ameloblasts in enamel²⁵, is time-dependent, such that new patterns can evolve as mineralization proceeds.

Inorganic morphosynthesis

Can an understanding of pattern formation in biomineralization be integrated within a synthetic approach to inorganic materials with complex form? Certainly the use of vesicles to expand the scale of materials patterning seems a plausible biomimetic strategy. Some important connections stand out. First, the biomineralized forms are generated by materials replication of patterned organic aggregates or supramolecular assemblies. Second, if the vesicles remain fixed in shape throughout mineralization, then the inorganic pattern is ‘transcribed’ from the self-assembled organic architecture. Alternatively, both mineral and associated vesicles develop synergistically, and new morphological patterns are added by coassembly of the inorganic and organic components. Third, continuous structures can be patterned around close-packed vesicle foams if the mineralization environment is restricted to the interstitial spaces and boundary edges.

The aim is to integrate these features into biomimetic strategies by using appropriate organic patterning agents to orchestrate the

TABLE 1 Pattern replication in inorganic morphosynthesis

Process	Pathway	Examples
Transcriptive synthesis	Self-assembly → Transcription → Replication	Tubular/lamellar/mesoporous silica, thin films, CdS arrays
Synergistic synthesis	Coadaptation → Coassembly → Replication	Mesoporous silica, metal oxides
Metamorphic reconstruction	Coassembly → Replication → Reconstruction	Microskeletal calcium phosphate, silica
Microphase separation	Coassembly → Evolution → Replication	Mesolamellar aluminophosphate, cellular calcium carbonate
Systems synthesis	Molecular assembly → Supramolecular assembly → Microphase assembly → System assembly	

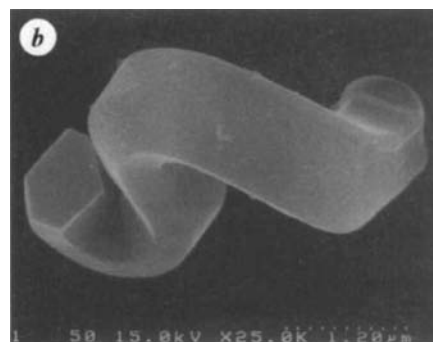
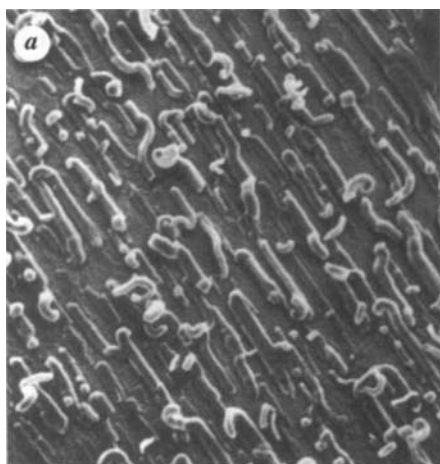


FIG. 3 a, As-synthesized mesoporous thin film grown on mica showing the initial nucleation of separate small oriented crystal ribbons with a preferred alignment⁴¹. (Field of view; width 49 μm , depth 51 μm .) b, Particle of mesoporous silica formed by bulk synthesis showing complex morphology⁴³. Dotted scale bar, 1.20 μm .

positioning, interconnection and stabilization of inorganic building blocks at the nano-, meso-, micro- and macroscopic level. We make the empirical assumption that the dimensionality of the patterning agent is commensurate with the length scale on which order and form are to be established. We highlight four mechanisms of inorganic morphosynthesis (Table 1). (1) Transcriptive synthesis involves stable, preorganized, self-assembled organic architectures serving as chemical and structural templates for patterned materials deposition. (2) Synergistic synthesis involves essentially one-step coassembly and pattern replication of inorganic precursors and associated organic molecules and aggregates, such as surfactants, micelles and microemulsions. (3) Metamorphic reconstruction involves coassembly and replication followed by *in situ* changes in the organization of the reaction field, so that new materials patterns develop. (4) In microphase separation, inorganic materials with microscopic and macroscopic patterning are formed as replicas of self-organized metastable structures, such as foams and vesicles.

Transcriptive synthesis. Implicit in this strategy is that the pattern and form of an inorganic material correspond closely with that of a preformed, self-assembled organic architecture. The template is preorganized and relatively stable, with chemical and morphological information 'written' into the surface structure (Table 1). Interfacial nucleation and crystal growth result in direct materials replication of the preformed organic shape. Several studies have investigated the use of compressed Langmuir monolayers^{26,27} and functionalized self-assembling monolayers^{28,29} as preorganized templates for the oriented nucleation of two-dimensional arrays of inorganic crystals and thin films. In many cases, the electrostatic, stereochemical and geometric properties of the surfactant headgroups are transcribed within the first layer of the inorganic nucleus by molecular recognition at the interface between monolayer and solution. This approach has been extended to more complex forms by using preformed, self-assembled bilayer templates that remain intact during and after inorganic mineralization. Such structures contain functionalized headgroups to facilitate site-directed nucleation as well as patterned architectures for materials replication. For example, certain biolipids have been used as preorganized fibrous templates for the production of hollow silica tubes³⁰, cylinders coated with iron oxide³¹, and helical strings of gold crystals³². Recently, lamellar silicas have been synthesized within the interlayer regions of multilamellar vesicle spheres³³. A similar approach has been used to produce coaxial cylinders of mesostructured aluminophosphates organized within the interlayers of multilamellar rod-shaped surfactant micelles³⁴.

Although the forms of these materials are reproducible, they are relatively simple, being based on spatially discrete organic architectures such as individual monolayer sheets, vesicle spheres,

lipid tubes and surfactant rods. More complex inorganic patterns require templates preorganized within extended phases, such as polymerized bicontinuous microemulsions, liquid crystals, polypeptide and polysaccharide networks, and organogels. For instance, liquid crystals of non-ionic surfactants have recently been used as direct templates for the synthesis of mesoporous silica³⁵ and cadmium sulphide superlattices³⁶. However, a common problem with this approach is that penetration of the inorganic precursors into the preformed organic materials is low, resulting in low mineral volume fractions and poor infiltration of the voids within the continuous framework, although this difficulty could be alleviated by incorporating artificial ion channels and auxiliary organics into the matrix.

Synergistic synthesis. We consider synergistic synthesis to imply that the development of mineralized patterns is based on cooperative interactions between inorganic and organic constituents present in the reaction media. A general feature of this process is the coadaptation of independent self-assembled systems with the formation of new organizational states replicated by materials deposition (Table 1). At the nanoscale, inorganic building blocks can be assembled around single organic molecules (silicate anions around quaternary ammonium cations in zeolite synthesis, for example³⁷). The association and ordering of the two components depends on the degree of chemical and structure complementarity at the inorganic-organic interface. The mechanism of zeolite formation has been proposed to involve the assembly of silicate and aluminate building blocks on curved, periodic minimal-energy surfaces having zeolite-like topologies³⁸. Such surfaces have their

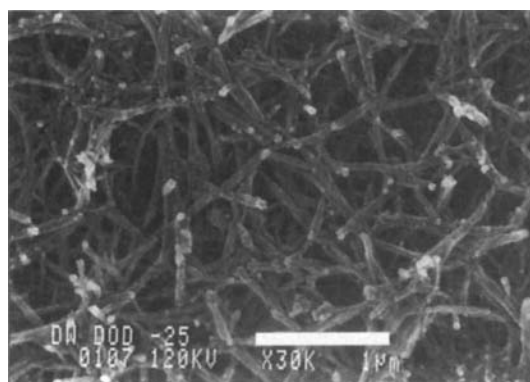


FIG. 4 Microskeletal framework of calcium phosphate formed by precipitation in bicontinuous microemulsions⁴⁶. Scale bar, 1 μm .

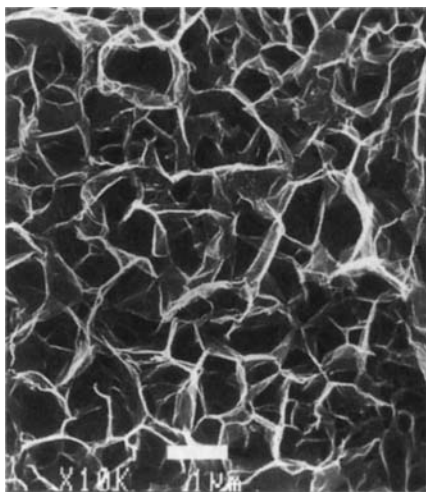


FIG. 5 SEM image of a thin cellular membrane of calcium carbonate formed by patterned mineralization of a biliquid oil/water foam⁴⁷. (Micrograph courtesy of D. Walsh, Univ. Bath.) Scale bar, 1 μm .

origin in density fluctuations of transient local order delineated by the organic template in a synthesis mixture.

These patterning processes can be extended to the mesoscale by using supramolecular aggregates of surfactant molecules^{2,39}. For example, the synthesis of mesoporous and mesolamellar silicas is initiated by ion binding and exchange of soluble silicate species at the cationic headgroups of quaternary ammonium surfactant molecules in weakly associated micellar and lamellar mesostructures, respectively⁴⁰. These cooperative interactions result in a change in the spatial charge density and steric requirements at the headgroup–silicate interface. The extent of coadaptation is very sensitive to the stoichiometry and relative chemical potentials of the reactants, that is, to the balance of thermodynamic and kinetic driving forces within the system. Thus a high supersaturation of silicate species induces phase separation through inorganic precipitation of amorphous silica, whereas low silicate concentrations give rise to soluble products. At intermediate concentrations, the interfacial energetics dominate so that lamellar, hexagonal or cubic arrays of the inorganic and organic constituents are coassembled either by *de novo* nucleation from ion pairs in solution, or by rearrangement and aggregation of coadapted micelles. At this stage, the assembled architecture is essentially a liquid-crystalline inorganic salt of the cationic surfactant, but can be further processed by ageing or increasing the temperature to induce the *in situ* condensation of the silicate species assembled within the organized biphasic array.

Two factors dominate the replication of the supramolecular assembly: the ability of oxyanions to undergo condensation by facile hydrolysis, and the flexibility in the bond angle of covalently linked Si–O–Si units. This ‘plasticity’ provides remarkable structural connectivity and stability, such that the 1-nm-thick curved walls enclosing the 3–5-nm cages remain intact after calcination of the organic template around 500 °C; the structural integrity is maintained until 1,100 °C. The resulting mesoporous silica is a kind of ‘crystalline glass’ with ordered mesopores whose inorganic walls are amorphous—an observation that highlights the hierarchical nature of these materials.

Recent studies indicate that continuous and oriented mesoporous, silica thin films can be synthesized on mica substrates⁴¹, presumably by the interaction of surfactant molecules with the mica surface which facilitates the organization and assembly of silica-surfactant micelles. This has been confirmed by atomic force microscopy studies of the surfactants adsorbed from aqueous solution onto mica, which reveal the presence of cylindrical aggregates rather than planar bilayers⁴². The silica mesophase

nucleates and grows into regular patterns of thin ribbons that are preferentially oriented within the plane of the mica substrate (Fig. 3a). Thus the synergistic assembly of the silica–surfactant biphasic is coupled with a transcriptive process involving direct templating at the mica surface. Similar interactions could be operating in the conventional synthesis of mesoporous silica in bulk solution. For example, particles of the biphasic material often exhibit a curved ‘worm-like’ morphology (Fig. 3b) that seems to be related to the topography of reaction spaces delineated by the higher-order organization of silica-surfactant vesicles in the synthesis gels⁴³.

Metamorphic reconstruction. It is possible that the synergistic synthesis of coassembled inorganic materials can be coupled interactively with the surrounding reaction medium. In principle, the nucleation and initial growth of an inorganic material within an organized multicomponent system, such as a water/surfactant/oil microemulsion or surfactant/water liquid crystal, can induce changes in the local structure and phase behaviour, such that new morphological patterns develop from existing architectures (Table 1). These changes might be relatively small and confined to localized reconstruction of the initial motif; for example, expansion of pore size (from 3 to 7 nm) in siliceous mesoporous materials by mild hydrothermal treatment of the surfactant biphasic products⁴⁴. Under alkaline conditions, part of the internal silica wall material is redistributed from regions of high to low curvature. The driving force for reorganization corresponds to a reduction in surface free energy, and occurs without loss of the imbibed surfactant template.

Alternatively, large changes in the scale and form of patterns replicated during mineralization could generate materials with complex architectures that do not resemble the initial organization of the associated reaction medium. For example, the mineralization reactions within the nanoscopic, interconnecting water channels of bicontinuous microemulsions produce microscopic, mesh-like architectures of silica⁴⁵ or calcium phosphate⁴⁶. The first stage in the formation of microskeletal calcium phosphate involves the coassembly of inorganic precursors (Ca^{2+} , HPO_4^{2-} , OH^-) and cationic surfactant molecules (didodecyldimethyl ammonium bromide (DDAB)) within an optically clear microemulsion made from the appropriate mixture of supersaturated aqueous solution, surfactant and tetradecane. Nucleation of hydroxyapatite ($\text{Ca}_{10}(\text{OH})_2(\text{PO}_4)_6$) at the surface of the DDAB headgroups, and subsequent growth of the crystals within the nanoscale water conduits for a few hours, results in filamentous networks of inorganic crystals with dimensions consistent with the microemulsion structure. But if samples are extracted after several days, an intact reticulated framework of micrometre-sized needle-like crystals, clearly incommensurate with the dimensions of the associated reaction media, is obtained (Fig. 4). The results suggest that the microskeletal inorganic form ‘evolves’ by localized disruption and reordering of the microemulsion due to incipient crystallization. Thus assembly of the inorganic framework seems to depend on the ability of the self-organized reaction

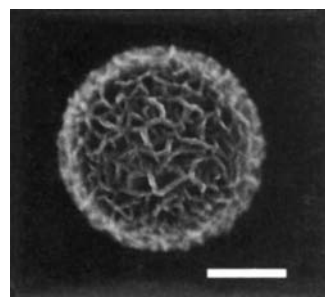


FIG. 6 Image of an intact 1- μm -diameter hollow shell of mesoporous aragonite⁴⁷.

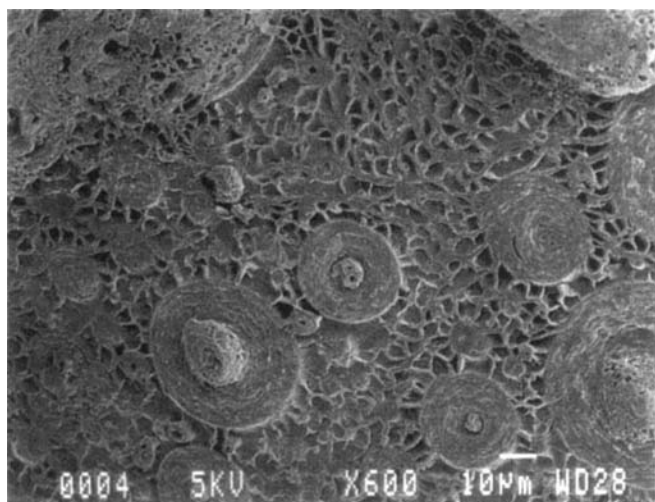


FIG. 7 Image showing complex surface patterns formed by mesolamellar aluminophosphate vesicles during hydrothermal synthesis. Scale bar, 10 μm .

environment to undergo restructuring or 'metamorphosis' during inorganic crystallization. Implicit in this transformability is that the surfactant headgroups initially act as the boundaries of nanoscale containers and templates but then become growth directors as crystallization proceeds.

In general, we consider that there could be metamorphic transformations running through various stages of hierarchical inorganic materials synthesis. Inorganic-organic composite materials are particularly susceptible because inefficient space-filling and mismatching in interfacial structure and charge could readily lead to metastability and *in situ* changes in form. Thus, patterns might become obscured by sequential deposition and subsequently hidden in the structural evolution of these materials.

Microphase separation. Extending the scale of inorganic morphosynthesis beyond the mesoscopic level necessitates the use of commensurate organic patterning agents. To this end, it seems feasible that the induction of localized microphase separation in multicomponent reaction systems—foams, vesicles and microemulsions—could be used to generate boundary surfaces and imprints for the sculpting of complex form in inorganic materials (Table 1). An important aspect of this strategy is that the formation of the microphase architectures must be synchronized with the onset of inorganic deposition, so that the self-organized patterns are replicated before disintegration.

One possibility is to use biliquid foams as patterned assemblies for synthesizing inorganic materials with honeycomb or mesh

morphologies. For example, solvent extraction of oil and surfactant from a supersaturated microemulsion film results in a thin cellular film of porous calcium carbonate (aragonite) (Fig. 5)⁴⁷. The microscopic mesh pattern is generated by demixing, which results in the evolution of close-packed oil droplets surrounded by a thin continuous layer of supersaturated aqueous calcium bicarbonate solution. This soapy foam is rapidly mineralized by loss of carbon dioxide to give an inorganic replica of the microphase-separated reaction media. Superficially, the system is analogous to the packing of areolar vesicles in the biomineralization of diatoms (see above)¹⁸. When uniform micrometre-sized polystyrene beads are coated in a thin film of the supersaturated microemulsion and washed with hot solvent, spherical shells of perforated calcium carbonate are formed. Dissolution and heat treatment of the polymer beads after mineralization results in intact hollow microspheres of the patterned inorganic material⁴⁷ (Fig. 6).

In principle, it should be possible to induce a series of localized phase separations on different length scales so that hierarchical patterns can be replicated. Recent studies involving the formation of a mesolamellar aluminophosphate by heat treatment of hydrated aluminium hydroxyoxide with a decylammonium dihydrogenphosphate lamellar precursor in tetraethylene glycol (TEG) have shown that a wide range of microscopic pores and surface patterns can be imprinted on unusual macroscopic morphologies such as solid spheroids and hollow shells (Fig. 7)⁴⁸. Some of these patterns display a striking resemblance to the biomineralized forms of diatoms and radiolaria, although we emphasize that mechanistically they are generated by different pathways. The macroscopic features, such as the bowl-shaped depressions and their associated mesh patterns, appear to be formed by the adhesion of mesolamellar aluminophosphate vesicles onto the surface of large spheres of the growing materials. An important aspect of this patterning process is that these higher-order vesicular aggregates can undergo microscopic phase separation together with fusion, fission, reshaping and collapse (Fig. 8). Replacement of TEG with ethylene glycol does not produce materials with complex forms, suggesting that the fine-scale architecture might arise from TEG-rich surface domains formed by phase separation within the vesicle bilayers⁴⁸. Overall, these observations illustrate how surfactants can be used to template a mesolamellar structure which in turn is self-patterned on longer length scales by multicomponent vesicles to produce a hierarchical organic-inorganic composite with complex morphology.

Future prospects

The development of inorganic morphosynthesis suggests that the seemingly disparate fields of biological assembly and inorganic materials chemistry can be conceptually linked. Where such connections have been made, for example in the chemical study of biomineralization and biomimetic materials chemistry⁵⁻¹¹, the level of integration has been formulated primarily from the viewpoint of the physical sciences. Perhaps it is time to take a 'biological view' of inorganic materials chemistry? For example, in this Article, biological concepts such as morphogenesis, replication, self-organization, transcription, synergism and metamorphosis have been used as constructs for thinking about novel ways of synthesizing inorganic materials exhibiting complex form across a range of length scales. Of course, a change of idiom is only significant if it leads to the design of new experiments and scientific objectives.

The use of hierarchical design concepts in inorganic morphosynthesis could lead to integrated systems with enhanced or highly adjustable mechanical properties, engineered biomaterials, bioactive ceramic-matrix composites, environmentally responsive systems, organized catalyst supports and more efficient membranes for liquid and gas separations, and water purification. In addition, understanding pattern replication across different length scales provides an opportunity to fabricate functional gradients in inorganic materials. We suggest that, in the longer

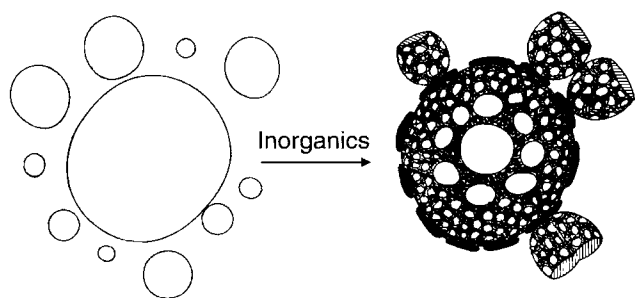


FIG. 8 Illustration of vesicle-templating of a hierarchical inorganic material. Mesolamellar aluminophosphate vesicles undergo fusion, fission, reshaping and collapse to form synthetic patterns with complex form⁴⁸.

term, the fabrication of bone and advanced materials will be described by the same paradigm of 'system-based synthesis' (Table 1). The constructional rules are likely to be different at different length scales, depending on the emerging properties and underlying patterns. Although the synthesis of such materials is far from being realized, the approaches highlighted in this

Article suggest that a new conceptual framework may soon become available. □

S.M. is at the School of Chemistry, University of Bath, Bath BA2 7AY, UK. G.A.O. is at the Materials Chemistry Research Group, Lash Miller Chemical Laboratories, University of Toronto, Toronto, Ontario M5S 1A1, Canada.

1. Barrer, R. M. *Hydrothermal Chemistry of Zeolites* (Academic, London, 1982).
2. Kresge, C. T., Leonowicz, M. E., Roth, W. J., Vartuli, J. C. & Beck, J. S. *Nature* **359**, 710–712 (1992).
3. Beck, J. S. et al. *J. Am. chem. Soc.* **114**, 10834–10843 (1992).
4. Mackay, A. L. *Curr. Sci.* **69**, 151–161 (1995).
5. Mann, S. *J. Mater. Chem.* **5**, 935–946 (1995).
6. Mann, S. *Nature* **332**, 119–124 (1988).
7. Heuer, A. H. et al. *Science* **225**, 1098–1105 (1992).
8. Mann, S. et al. *Science* **261**, 1286–1292 (1993).
9. Mann, S. *Nature* **365**, 499–505 (1993).
10. Ozin, G. A. & Oliver, S. *Adv. Mater.* **7**, 943–947 (1995).
11. Oliver, S., Ozin, G. A., Khushalani, D., Kuperman, A. & Coombs, N. in *NATO Adv. Res. Workshop Proc.: Modular Chemistry* (ed. Michl, J.) (Aspen, Colorado, 1995).
12. Garcia-Ruiz, J. M. *J. Cryst. Growth* **73**, 251–262 (1985).
13. Kumar, A., Abbott, N. L., Kim, E., Biebuyck, H. A. & Whitesides, G. *Acc. Chem. Res.* **28**, 219–226 (1995).
14. Wilbur, J. L., Kim, E., Xia, Y. & Whitesides, G. *Adv. Mater.* **7**, 649–652 (1995).
15. Harting, P. Q. *J. Microsc. Sci.* **12**, 118–123 (1872).
16. Thompson, D. W. *On Growth and Form* (Cambridge Univ. Press, 1942).
17. Anderson, O. R. in *Biomaterialization in Lower Plants and Animals* (eds Leadbeater, B. S. C. & Riding, R.) 375–391 (Systematics Assocn Vol. 30, Oxford Univ. Press, 1986).
18. Crawford, R. M. & Schmid, A. M. M. in *Biomaterialization in Lower Plants and Animals* (eds Leadbeater, B. S. C. & Riding, R.) 290–314 (Systematics Assocn Vol. 30, Oxford Univ. Press, 1986).
19. Leadbeater, B. S. C. *Proc. R. Soc. Lond.* **B304**, 529–536 (1984).
20. Westbroek, P., van der Wal, P., van Emburg, P. R., de Vrind-de Jong, E. W. & de Bruijn, W. C. in *Biomaterialization in Lower Plants and Animals* (eds Leadbeater, B. S. C. & Riding, R.) 189–203 (Systematics Assocn Vol. 30, Oxford Univ. Press, 1986).
21. Ledger, P. W. & Jones, W. C. *Cell Tiss. Res.* **181**, 553–567 (1977).
22. Märkel, K., Röser, U., Mackenstedt, U. & Klosterman, M. *Zoomorphology* **106**, 232–243 (1986).
23. Wilbur, K. M. & Saleuddin, A. S. M. in *The Mollusca* Vol. 4 (eds Saleuddin, A. S. M. & Wilbur, K. M.) 235–287 (Academic, New York, 1983).
24. Currey, J. *The Mechanical Adaptations of Bones* (Princeton Univ. Press, New Jersey, 1984).
25. Weiner, S. *CRC Crit. Rev. Biochem.* **20**, 365–408 (1986).
26. Heywood, B. R. & Mann, S. *Adv. Mater.* **6**, 9–20 (1994).
27. Fendler, J. H. & Meldrum, F. C. *Adv. Mater.* **7**, 607–632 (1995).
28. Feng, S. & Bein, T. *Nature* **368**, 834–836 (1994).
29. Shin, H., Collins, R. J., de Guire, M. R., Heuer, A. H. & Sukenik, C. N. *J. Mater. Res.* **10**, 692–703 (1995).
30. Baral, S. & Schoen, P. *Chem. Mater.* **5**, 145–147 (1993).
31. Archibald, D. D. & Mann, S. *Nature* **364**, 430–433 (1993).
32. Burkett, S. L. & Mann, S. *Chem. Commun.* 321–322 (1996).
33. Taney, P. T. & Pinnavaia, T. J. *Science* **271**, 1267–1269 (1996).
34. Chenite, A., Le Page, Y., Karra, V. R. & Sayari, A. *Chem. Commun.* 411–413 (1996).
35. Attard, G. S., Glyde, J. C. & Goltner, C. G. *Nature* **378**, 366–368 (1995).
36. Braun, P. V., Osenar, P. & Stupp, S. I. *Nature* **380**, 325–328 (1996).
37. Zones, S. I. & Davis, M. E. *Current Opin. Solid St. Mater. Sci.* **1**, 107–117 (1996).
38. Anderson, S., Hyde, S. T., Larsson, K. & Lidin, S. *Chem. Rev.* **88**, 221–240 (1988).
39. Huo, Q. et al. *Chem. Mater.* **6**, 1176–1191 (1994).
40. Firoouzi, A. et al. *Science* **267**, 1138–1143 (1995).
41. Yang, H., Coombs, N., Kuperman, A. & Ozin, G. A. *Nature* **379**, 703–705 (1996).
42. Manne, S. & Gaub, H. *Science* **270**, 1480–1482 (1995).
43. Khushalani, D., Kuperman, A., Oliver, S. & Ozin, G. A. in *NATO Adv. Res. Workshop. Proc.: Modular Chemistry* (ed. Michl, J.) (Aspen, Colorado, 1995).
44. Khushalani, D., Coombs, N., Kuperman, A. & Ozin, G. A. *Adv. Mater.* **7**, 842–847 (1995).
45. Watzke, H. J. & Dieschbourg, C. *Adv. Colloid Interface Sci.* **50**, 1–14 (1994).
46. Walsh, D., Hopwood, J. D. & Mann, S. *Science* **264**, 1576–1578 (1994).
47. Walsh, D. & Mann, S. *Nature* **377**, 320–323 (1995).
48. Oliver, S., Kuperman, A., Coombs, N., Lough, A. & Ozin, G. A. *Nature* **378**, 47–50 (1995).
49. Phil. *Trans. R. Soc. Lond.* **B 304**, 518 (1984).

CORRESPONDENCE should be addressed to S.M. (e-mail: s.mann@bath.ac.uk) or G.A.O. (email: gozin@alchemy.chem.utoronto.ca).

KNOW YOUR COPY RIGHTS RESPECT OURS

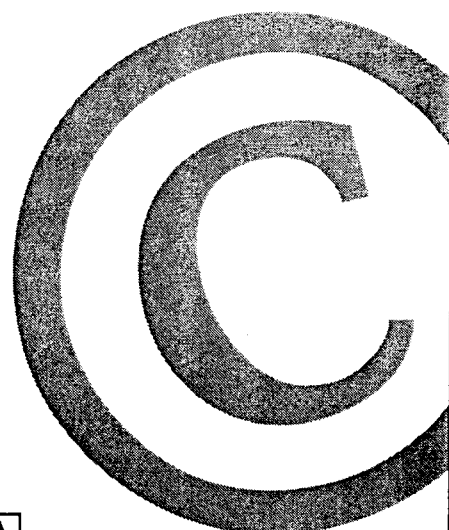
The publication you are reading is protected by copyright law. This means that the publisher could take you and your employer to court and claim heavy legal damages if you make unauthorised photocopies from these pages.

Photocopying copyright material without permission is no different from stealing a magazine from a newsagent, only it doesn't seem like theft.

The Copyright Licensing Agency (CLA) is an organisation which issues licences to bring photocopying within the law. It has designed licensing services to cover all kinds of special needs in business, education, and government.

If you take photocopies from books, magazines and periodicals at work your employer should be licensed with CLA.

Make sure you are protected by a photocopying licence.



The Copyright Licensing Agency Limited
90 Tottenham Court Road, London W1P 0LP
Telephone: 0171 436 5931
Fax: 0171 436 3986

A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A

Xiang-Jiao Yang, Vasily V. Ogryzko, Jun-ichi Nishikawa*, Bruce H. Howard & Yoshihiro Nakatani

Laboratory of Molecular Growth Regulation, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892-2753, USA

The adenoviral oncoprotein E1A induces progression through the cell cycle by binding to the products of the p300/CBP and retinoblastoma gene families. A new cellular p300/CBP-associated factor (P/CAF) having intrinsic histone acetylase activity has been identified that competes with E1A. Exogenous expression of P/CAF in HeLa cells inhibits cell-cycle progression and counteracts the mitogenic activity of E1A. E1A disturbs the normal cellular interaction between p300/CBP and its associated histone acetylase.

THE transforming proteins encoded by adenovirus and several other small DNA tumour viruses disturb host cell-growth control by interacting with cellular factors that normally function to repress cell proliferation. One of the most intensively studied of these viral proteins, the product of the adenovirus E1A gene, is itself sufficient for transformation¹. E1A transforming activity resides in two distinct domains, the targets of which include p300 and products of the retinoblastoma (RB)-susceptibility gene family^{1,2}. Interactions of E1A with p300 and RB are thought to influence functionally distinct growth-regulatory pathways, allowing the two domains to contribute additively to transformation¹.

The way in which E1A and functionally related viral proteins perturb cell-growth regulation is understood in large part from studies of their interactions with RB protein^{1,2}. The molecular function of E1A is based on its capacity to interfere with cellular protein-protein interactions. As both E1A and various cellular targets bind to a site in RB termed the pocket domain², E1A can competitively disrupt complex formation between RB and its cellular targets.

The second cellular factor implicated in E1A-dependent transformation, p300, is believed to inhibit exit from the G0/G1 phases of the cell cycle, to activate certain enhancers, and to stimulate differentiation^{1,2}. p300 is very closely related to CBP, a factor that binds selectively to the protein-kinase-A-phosphorylated form of CREB³⁻⁵. p300 and CBP exhibit extensive amino-acid sequence similarity and share the capacity to bind both CREB and E1A (refs 6-8). Although neither p300 nor CBP by itself binds to DNA, each can be recruited to promoter elements by interaction with sequence-specific activators. These interactions include p300-CREB (ref. 7), CBP-CREB (refs 3,4), CBP-c-Jun (ref. 9), p300-YY1 (ref. 10), CBP-Fos (ref. 11), CBP-c-Myb (ref. 12) and CBP with nuclear receptors¹³. Although some interactions have been reported for either p300 or CBP only, there is no evidence to indicate that there is a specificity between p300 and CBP in binding with activators. For simplicity, p300 and CBP will be termed p300/CBP when discussing their shared functional properties.

E1A inhibits the p300/CBP-mediated transcriptional activation of many promoters (see ref. 14 and refs therein). In one case, the complex of p300 and YY1, E1A inhibits transcription without

disrupting the complex¹⁰, indicating that a cellular counterpart of E1A may be present that binds to p300/CBP and is important in both transcription and cell-cycle regulation. We now present the cloning and characterization of a new p300/CBP-associated factor (P/CAF) with intrinsic histone acetylase activity. Both E1A and P/CAF bind to the same or to very closely spaced sites on p300/CBP and compete *in vitro*. The *in vivo* interaction is revealed by co-immunoprecipitation of P/CAF and p300/CBP from cell extracts. The endogenous P/CAF-p300 complex is disrupted by E1A *in vivo*. Intriguingly, exogenous expression of P/CAF in HeLa cells inhibits cell-cycle progression and counteracts the mitogenic activity of E1A. These results are consistent with the view that E1A targets the interaction between P/CAF and p300/CBP in inducing cell-cycle progression.

Cloning of P/CAF

We cloned the p300/CBP-associated factor (P/CAF) on the basis of an analogy between species. In human cells, CBP binds to c-Jun in a phosphorylation-dependent manner in association with stimulation of transcription⁹. In yeast, GCN4 is believed to be a c-Jun counterpart on the basis of similarities in DNA recognition¹⁵, as well as the fact that both proteins participate in ultraviolet signalling pathways¹⁶. Yeast genetic screening has led to the isolation of various cofactors for GCN4, including GCN5 (yGCN5), ADA2 (yADA2) and ADA3 (yADA3)¹⁷⁻¹⁹. These factors are considered to function as a complex (or in a common pathway) on the basis of genetic and protein-protein interaction studies¹⁸⁻²². Finally, p300/CBP and yADA2 share significant sequence similarity within a 50-amino-acid region that includes a zinc-finger motif³. From these observations, we considered that human counterparts might exist to yGCN5, yADA2 or yADA3 which could interact with p300/CBP to mediate transcriptional activation by c-Jun.

Comparison of the yGCN5 protein sequence with various databases²³ revealed significant similarities with randomly sequenced human complementary DNAs, EST05039 (ref. 24) ($P = 4.0 \times 10^{-15}$) and NIB2000-5R ($P = 6.5 \times 10^{-9}$). Given that these cDNAs were truncated, complete clones were isolated from human cDNA libraries. The complete sequences revealed that the ETS05039- and NIB2000-5R-derived clones are encoded by distinct genes but are highly related within the protein-coding regions (68% identity at the DNA level; 75% identity and 86% similarity at the protein level). The former encodes an N-terminal region

* Present address: Department of Environmental Biochemistry, Faculty of Pharmaceutical Sciences, Osaka University, 1-6 Yamada-oka, Suita, Osaka 565, Japan.

with no obvious sequence similarity to any proteins in the databases besides the yGCN5-related C-terminal region, whereas the latter encodes only the yGCN5-related region (Fig. 1a, b). Given that the former polypeptide had obvious p300/CBP-binding activity (see below), we named it p300/CBP-associated factor (P/CAF) and the latter human GCN5 (hGCN5).

RNA blotting indicates that transcripts detected by the P/CAF (Fig. 1c) and hGCN5 (data not shown) cDNAs are ubiquitously expressed, but the former is most abundant in heart and skeletal muscle, whereas the latter is most abundant in pancreas and skeletal muscle.

P/CAF-p300/CBP interaction *in vitro*

We considered that one third of the C-terminus of CBP (residues 1,678–2,442) may contain the P/CAF-binding site as this region fused to a DNA-binding domain activates transcription⁴ in a manner repressed by coexpression of 12S E1A (X.-J.Y. and Y.N., unpublished observations). This region was divided into six overlapping fragments (Fig. 2a), and each was expressed in *Escherichia coli* as a glutathione-S-transferase (GST) fusion protein. GST-CBP fusions were incubated with recombinant P/CAF protein and subsequently purified using glutathione-Sepharose. Co-purified P/CAF was detected by immunoblotting (Fig. 2b); two CBP segments, A and B, interacted specifically with P/CAF (lanes 3 and 4). The stronger interaction was observed in the latter segment, which does not include the yADA2-like zinc-finger. Given that the CBP segment B is well conserved in p300 (66% identity, 75% similarity), we tested whether P/CAF also binds to p300 *in vitro*. For this experiment, the p300 segment spanning residues 1,763–1,966, termed segment B', which is analogous to the CBP segment B, was used. Like CBP, p300 interacted specifically with P/CAF (Fig. 2b, lane 9). We conclude that P/CAF binds specifically to both p300 and CBP *in vitro*. In contrast to P/CAF, hGCN5 did not bind to CBP or p300 (Fig. 2c).

Previous reports indicated that E1A binds to both the p300 segment spanning residues 1,767–1,816 and the CBP segment spanning residues 1,805–1,854 (ref. 7). These interactions were reconfirmed in our system; thus, both p300 and CBP segments covering the previously identified regions interacted with E1A (Fig. 2d, lanes 3, 4 and 9).

For further mapping, a series of deletions was introduced within the CBP segment B and tested for interaction with P/CAF (Fig. 2Ba) and E1A (Fig. 2Bb). Deletions of residues 1,801–1,825 or 1,824–1,851 markedly reduced interactions with both P/CAF and E1A (Fig. 2Ba, b; lanes 3 and 4), whereas deletion of residues 1,850–1,878 did not affect these interactions (lanes 5). Furthermore, deletion of residues 1,801–1,851 completely abolished interactions with both P/CAF and E1A (lanes 6). This indicates that residues 1,801–1,851 of CBP are critical for interaction with both P/CAF and E1A. Taken together with the evidence that CBP segment A (residues 1,678–1,880) also binds to these factors (Fig. 2Ab, d), our findings

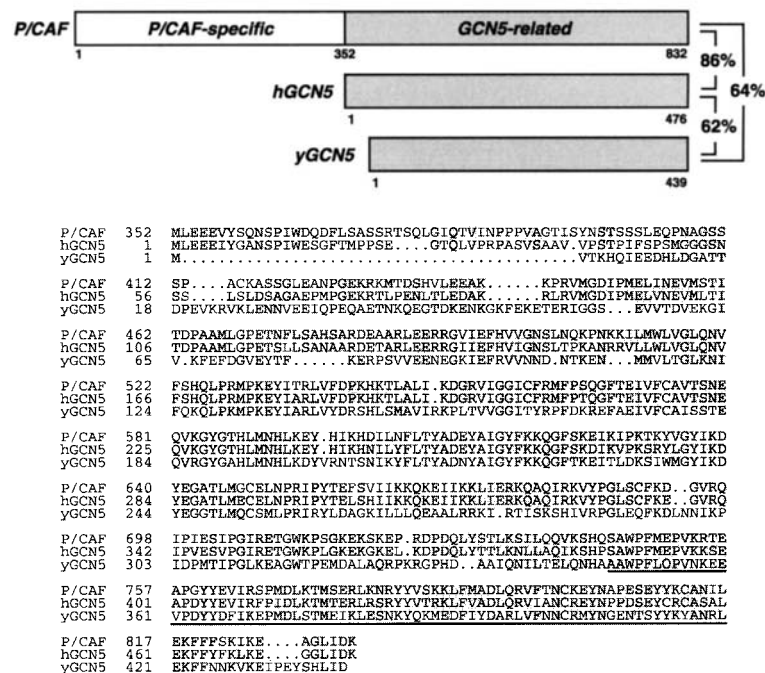
a

```

1  MSEAGGAGPG  GCGAGAGAGA  GPGALPPQPA  ALPPAPPQGS  PCAAAAGSGS
51  ACGPATAVAA  AGTAEGPGGG  GSARIAVKA  QLRSAAPRAK  LEKLVVYSAC
101 KAESCKCKNG  WKNPNPSPPT  PRADLQIIV  SLTESCSCSC  HALAAHVSHL
151 ENVSEEMNMR  LLGIVLDVEY  LFTCVHKEED  ADTKQVYFVL  FKLRLKSLIQ
201 RGKPVVEGSL  EKKPPEFEKS  IEQGVNPFVQ  YKFSHLPAGE  RQTIVELAKM
251 FLNRIYVHSL  EAPSQRRLRS  PNDDISGYKE  NYTRWLCYCN  VPQGCDSLPR
301 YETTVQVGR  LLRSVFTVMR  RQLLEQARQE  KDKLPLEKRT  LLLTHFPKFL
351 SMLEEVVYSQ  NSPIWDQDFL  SASSRTSQLG  IQTVINPPPV  AGTISYNSYS
401 SSLEQPNAGS  SSPACKASSG  LEANPGEKRR  MTDHSHVLEA  KPRVVMGDI
451 MELINEVMST  ITDPAAMLGP  ETNFLSAHSA  RDEAARLEER  RGVIEFHVVG
501 NSLNQKPNKK  ILMWLVLQGN  VFSHQLPRMP  KEYITRLVFD  PKHKTALIK
551 DGRVIGGICF  RMFSPQGFT  IVFCAVTSNE  QVKGYGTHLM  NHLKEYHIKH
601 DILNFLTAYD  EYAIQYFKKQ  GFSKEIKIPK  TKYVGYIKDY  EGATLMGCEL
651 NRPITYTEFS  VIKKQKEIK  KLIERKQAO  IRKVYVGLSC  FKDGVRQIPI
701 ESIPGIRETG  WKPSGKEKSK  EPRDPQLYS  TLKSLQVQVK  SHQSAWPFME
751 PVKRTAPGY  YEVIRSPMDL  KTMSERLKNR  YVVSCKLFMA  DLQRVFTNCK
801 EYNAPSEYY  KCANILEKFF  FSKIEAGLI  DK

```

b



c

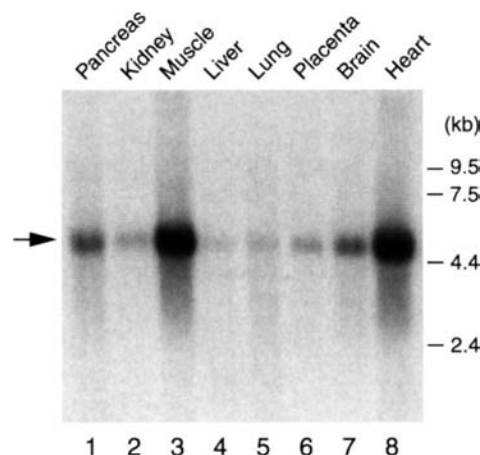


FIG. 1 a, Deduced amino-acid sequence of P/CAF. b, Comparison of the deduced amino-acid sequence of P/CAF, hGCN5 and yGCN5 (ref. 17). P/CAF has the non-conserved N-terminal (unshaded) and the GCN5-related C-terminal (shaded) regions (top). Sequence similarities of the conserved C-terminal regions are indicated. Sequence alignment of the C-terminal conserved region (bottom) is shown. Identical and related amino acids are shaded. The position of a potential structural motif, the bromodomain²⁹, is underlined. Note that our hGCN5 has an additional 49 amino acids at the N terminus compared with the published sequence²⁵. c, Expression profile of P/CAF in various human tissues.

METHODS. a, Human fetal liver and fetal brain cDNA libraries (Clontech) were screened with ETS05039 and NIB2000-5R, respectively. c, An RNA blot (Clontech) was hybridized with a random-primed probe made from the cDNA encoding P/CAF.

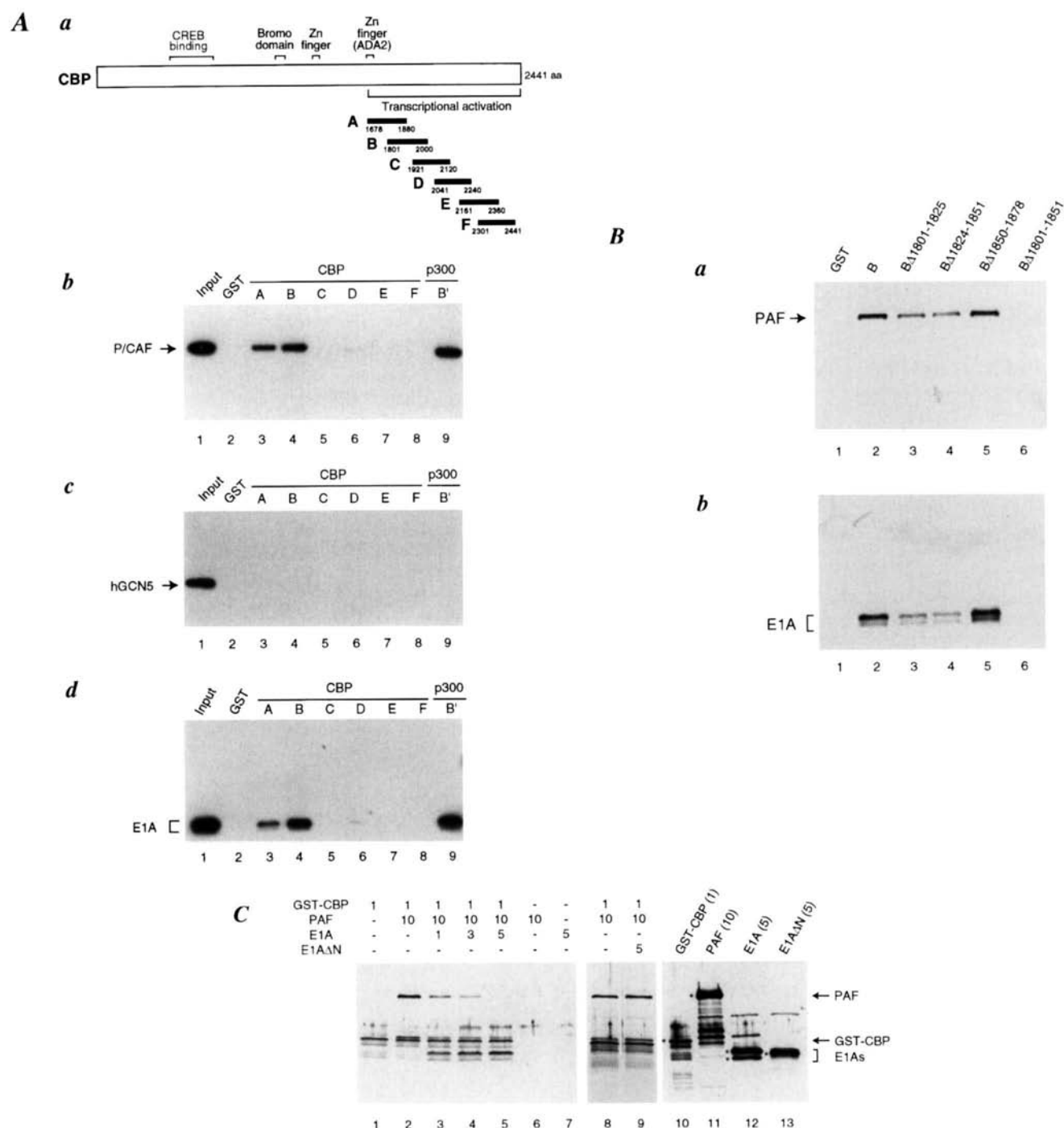


FIG. 2 A, P/CAF-p300/CBP interaction *in vitro*. **a**, Positions of the CBP segments used for *in vitro* interaction experiments. For p300 interactions, the segment spanning residues 1,763–1,966 (segment B') of p300, which is analogous to the CBP segment B, was used. **b–d**, Interactions of CBP (lanes 3–8) or p300 (lanes 9) segments with P/CAF (**b**), hGCN5 (**c**) or E1A (**d**). Twenty per cent of the P/CAF and hGCN5 inputs and all of the E1A input were also analysed (**b–d**, lanes 1). **B**, P/CAF (**a**) and E1A (**b**) binding sites within the CBP segment B. In the GST precipitation assays in **A** and **B**, almost identical amounts of the GST fusions were recovered in all samples (data not shown). **C**, Inhibition of the P/CAF-CBP interaction by E1A *in vitro*. Interaction between P/CAF and CBP (segment B) was determined in the absence (lanes 2 and 8) and in the presence of wild-type (lanes 3–5) or mutant (lane 9) E1A. Control reactions with GST–CBP alone (lane 1) and without GST–CBP (lanes 6 and 7) were also done. Input proteins were analysed (lanes 10–13). The bands corresponding to GST–CBP, P/CAF, and E1A are indicated by arrows and asterisks. The molar ratios of proteins are indicated above each lane.

METHODS. To construct GST fusions, the indicated regions of CBP and

p300 were amplified by PCR. A series of deletions of the CBP segment B was created by site-directed *in vitro* mutagenesis. These fragments were subcloned into pGEX-2T (Pharmacia). GST fusions were expressed in *E. coli*, extracted with buffer B (20 mM Tris-HCl, pH 8.0, 5 mM MgCl₂, 10% glycerol, 1 mM 4-(2-Aminoethyl)benzenesulphonyl fluoride (AEBSF), 0.1% NP-40, 10 μg ml⁻¹ aprotinin, 10 μg ml⁻¹ leupeptin, 1 μg ml⁻¹ pepstatin A, 1 mM DTT) containing 0.1 M KCl and used for **A** and **B**. GST–CBP-segment B was purified on glutathione–Sepharose and phenyl–Sepharose for **C**; P/CAF, hGCN5 and E1A were expressed as FLAG fusions in Sf9 cells via baculovirus vectors and affinity-purified with M2 agarose³⁰ (Kodak-IBI). For interaction, a crude *E. coli* extract containing 20 pmol GST fusion protein was incubated with 40–60 pmol of P/CAF or E1A in a total volume of 50 μl buffer B with 0.1 M KCl on ice for 10 min. Samples were further incubated with 10 μl (packed volume) of glutathione–Sepharose at 4 °C for 30 min, washed four times with 200 μl buffer B containing 0.1 M KCl, and eluted with 20 μl buffer E (50 mM Tris-HCl, pH 8.0, 0.2 M KCl, 20 mM glutathione) for 60 min. Interacting proteins were detected by anti-FLAG immunoblotting (**A** and **B**) or silver staining (**C**).

lead us to conclude that P/CAF and E1A bind to the same, or very closely spaced, sites on CBP.

In contradiction to our original idea, the zinc-finger region of p300/CBP, which shares sequence similarity with yADA2, is not essential for the interaction with P/CAF (Fig. 2A). In related work, we and others have also cloned a human structural homologue of yADA2, termed hADA2 (ref. 25; X.-J.Y. and Y.N., unpublished results). In contrast to the sequence similarity between p300/CBP and yADA2, which is restricted to a 50-amino-acid region, hADA2 shares extensive similarity (30% identity, 52% similarity) with yADA2 over the entire protein sequence. Moreover, a computer search of the complete genomic sequence of *Saccharomyces cerevisiae* revealed that yeast has no obvious counterparts of p300/CBP or P/CAF. Thus, the p300/CBP-P/CAF pathway may have been acquired during metazoan evolution.

Action of E1A *in vitro*

Evidence that both P/CAF and E1A recognize the same p300/CBP segments suggests that P/CAF and E1A compete directly for binding to p300/CBP. To test this possibility, we used affinity-purified recombinant proteins in a competition experiment (Fig. 2C). The interaction of P/CAF with CBP segment B (lane 2) was progressively inhibited by addition of increasing amounts of E1A (lanes 3–5). No inhibition was caused by an E1A mutant that does not bind p300/CBP (E1AΔN) (lanes 8 and 9). Similar results were obtained with p300 segment B' (data not shown). We conclude that P/CAF and E1A compete for the binding sites in p300/CBP.

P/CAF-p300/CBP interaction *in vivo*

The *in vivo* interaction between P/CAF and p300/CBP was established by co-immunoprecipitation from a human osteo-

sarcoma cell extract (Fig. 3A). Proteins in this extract were immunoprecipitated with anti-P/CAF antibody or control IgG, and the precipitates analysed by immunoblotting. Anti-P/CAF antibody specifically detected a protein of relative molecular mass 95,000 (*M*, 95K), which is very close to the calculated value for full-length P/CAF, in the immunoprecipitates (lane 1). Significantly, anti-P/CAF antibody co-immunoprecipitated both CBP (lane 4) and p300 (lane 7). Similarly, anti-CBP antibody also co-immunoprecipitated P/CAF (lane 2). However, anti-p300 antibody did not co-immunoprecipitate P/CAF (data not shown). This is probably due to steric interference because the anti-p300 antibody was raised against the p300 segment spanning residues 1,572–2,371 (ref. 5), which includes the P/CAF-binding region. Thus, we conclude that P/CAF forms complexes with both p300 and CBP *in vivo*.

Action of E1A *in vivo*

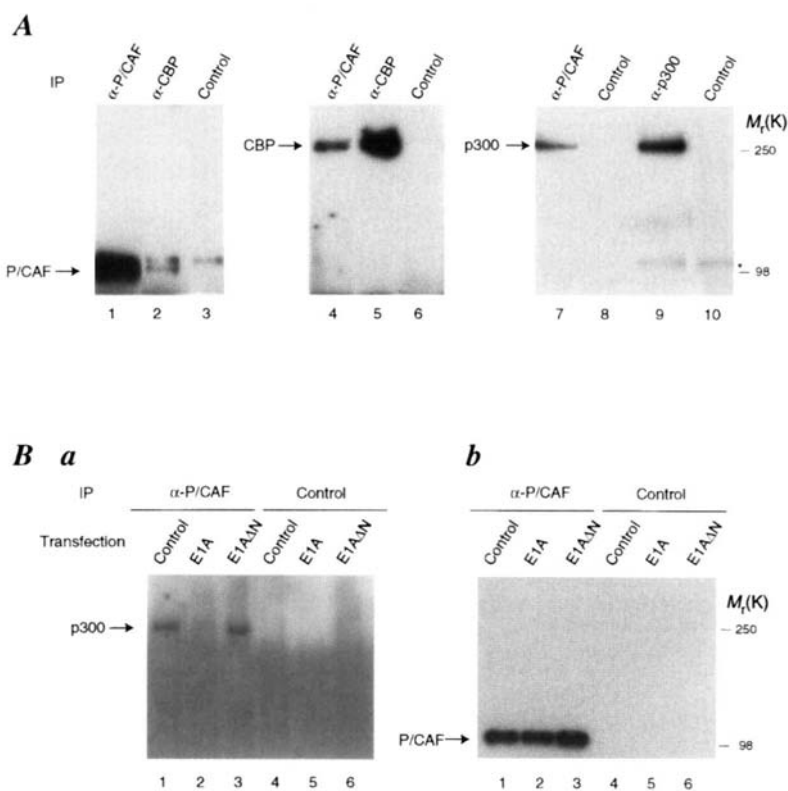
In vitro experiments indicate that P/CAF and E1A compete for binding sites in p300/CBP (Fig. 2). We tested whether E1A targets the endogenous interaction between P/CAF and p300 (Fig. 3B). An E1A-expression vector was transiently transfected into human osteosarcoma cells, and the transfected subpopulation purified by cell sorting. The interaction between P/CAF and p300 in transfected cells was then examined by co-immunoprecipitation with anti-P/CAF antibody. The endogenous interaction of P/CAF with p300 (Fig. 3Ba, b, lanes 1) was drastically inhibited by expression of E1A (lanes 2). No inhibition was observed for the E1A mutant lacking the p300-binding domain (E1AΔN) (lanes 3). We conclude that E1A disrupts the P/CAF-p300 complex *in vivo* through interaction with p300.

Cell-cycle regulation by P/CAF

Given that binding of P/CAF to p300/CBP is inhibited by E1A, we

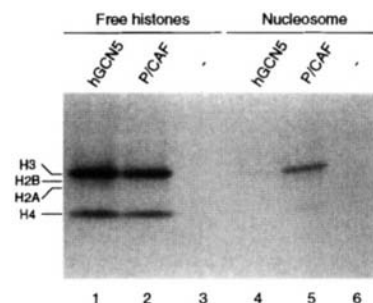
FIG. 3 A, P/CAF-p300/CBP interaction *in vivo*. Cell extract was immunoprecipitated with rabbit anti-P/CAF (lanes 1, 4 and 7), rabbit anti-CBP (lanes 2 and 5), and mouse anti-p300 (lane 9) antibodies. For controls, cell extract was precipitated with rabbit control IgG (lanes 3, 6 and 8) or mouse anti-HA monoclonal antibody (lane 10). The precipitates were analysed by immunoblotting with anti-P/CAF (lanes 1–3), anti-CBP (lanes 4–6), and anti-p300 (lanes 7–10) antibodies. The positions of non-specific bands are indicated by asterisks. B, E1A inhibits the P/CAF-p300 interaction *in vivo*. Osteosarcoma cells were transfected with either control vector (lanes 1 and 4) or E1A (lanes 2 and 5) or E1AΔN (lanes 3 and 6) expression vectors. Extract from the transfected subpopulation was immunoprecipitated with anti-P/CAF (lanes 1–3) or control (lanes 4–6) IgG. Precipitates were analysed by immunoblotting with anti-p300 (a) and anti-P/CAF (b).

METHODS. A, Rabbit anti-P/CAF antibody was raised against the P/CAF segment spanning residues 125–397 and purified by immunoaffinity chromatography³¹. A mixture of monoclonal antibodies raised against the human p300 segment spanning residues 1,572–2,371 (ref. 5) and rabbit polyclonal antibodies raised against the mouse CBP segment spanning residues 2–23 (for immunoprecipitation) and 1,736–2,179 (immunoblotting) were from Upstate Biotechnology. Approximately 2×10^7 human osteosarcoma U2OS cells were extracted with 10 ml lysis buffer (25 mM HEPES-KOH, pH 7.2, 150 mM potassium acetate, 2 mM EDTA, 1 mM DTT, 1 mM AEBSF, $10 \mu\text{g ml}^{-1}$ aprotinin, $10 \mu\text{g ml}^{-1}$ leupeptin, $1 \mu\text{g ml}^{-1}$ pepstatin A, 20 mM sodium fluoride, 0.1% NP-40). 2–10 ml of extract were incubated with 2 μg of the respective antibody for 4 h at 4 °C. 50 μl (packed volume) of protein A-Trisacryl (Pierce) were added, and incubation continued for 2 h. The matrix was washed four times with 1 ml lysis buffer, then boiled in $2 \times$ SDS sample buffer. B, Human osteosarcoma



U2OS cells were transfected with 20 μg of indicated plasmid and 1 μg sorting plasmid (pCMV-IL2R)³². The transfected subpopulation was purified by magnetic affinity cell sorting³³. Extract from $\sim 2 \times 10^5$ sorted cells was immunoprecipitated as described for A.

FIG. 5 Histone acetylase activity of P/CAF. Activity of hGCN5 (lanes 1 and 4) and P/CAF (lanes 2 and 5) which acetylates free histones (lanes 1–3) or histones in the nucleosome core particle³⁵ (lanes 4–6) was measured as described³⁶. Each reaction contained 0.3 pmol affinity-purified FLAG-hGCN5 or FLAG-P/CAF, 4 pmol histone octamer or nucleosome core particle, and 10 pmol [14 C]acetyl-CoA. Note that the histone octamer dissociates into dimers or tetramers under assay conditions. Acetylated histones were detected by autoradiography after separation by SDS-PAGE. Bands corresponding to acetylated histones H3 and H4 are indicated.



considered that P/CAF, by binding to and forming a functional complex with p300, might be involved in the regulation of entry into S phase. This possibility was addressed by examining whether transient expression of P/CAF would affect the rate of G1/S transit in HeLa cells. P/CAF negatively affected the distribution of cells between G1 and S phases in this assay. The fraction of cells accumulating in S phase in control cultures was 23%, compared to 15% in P/CAF-transfected cells (Fig. 4a, b, respectively). This effect was reproducible in several independent experiments. In parallel experiments to test the validity of this experimental

protocol, plasmids encoding E2F-1, simian virus 40 small-t antigen, cyclin A or cyclin E increased the accumulation of cells in S phase, whereas plasmids encoding the cyclin-dependent-kinase inhibitors p21 or p27 reduced the number of S-phase cells (data not shown).

Based on evidence that E1A and P/CAF compete for binding sites on p300, it might be expected that cotransfection of P/CAF with E1A would oppose the mitogenic effect caused by E1A. E1A alone has mitogenic activity under these conditions (Fig. 4c), whereas the E1A mutant lacking the p300-binding domain

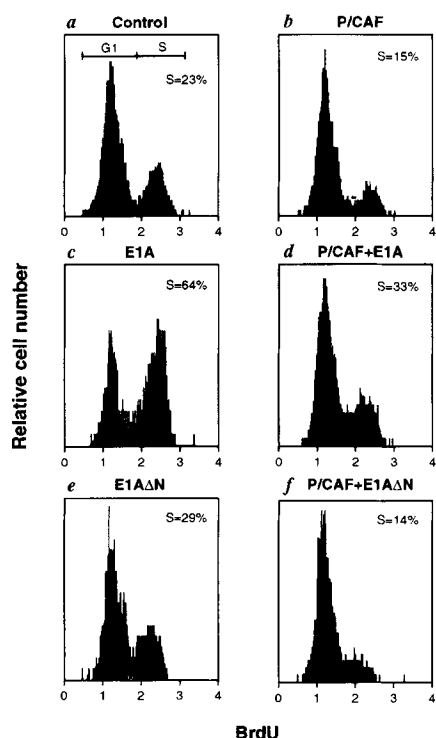


FIG. 4 P/CAF and E1A mediate antagonistic effects on cell-cycle progression. HeLa cells were transfected by electroporation with 7 μ g P/CAF expression plasmid and/or 3 μ g of full-length or N-terminally deleted (Δ 2–36) E1A 12S expression plasmid as indicated. Plasmids were constructed by subcloning FLAG-P/CAF and E1A cDNAs into pCX (ref. 34) and pcDNA1 (Invitrogen), respectively. All samples in addition contained 1 μ g sorting plasmid (pCMV-IL2R)³² and carrier plasmid (pCX) to normalize the total amount of DNA to 11 μ g. After transfection, cells were incubated in Dulbecco's modified Eagle's medium with 10% fetal bovine calf serum for 12 h, and subsequently labelled in medium containing 10 μ M bromodeoxyuridine (BrdU) for 30 min. Subsequently, the transfected subpopulation was purified by magnetic affinity cell sorting, and nuclei were analysed by dual parameter flow cytometry as described³³. Histograms show percentages of cells in G1 and S phases. Abscissa values represents fluorescence intensity of bound anti-BrdU antibodies in log scale.

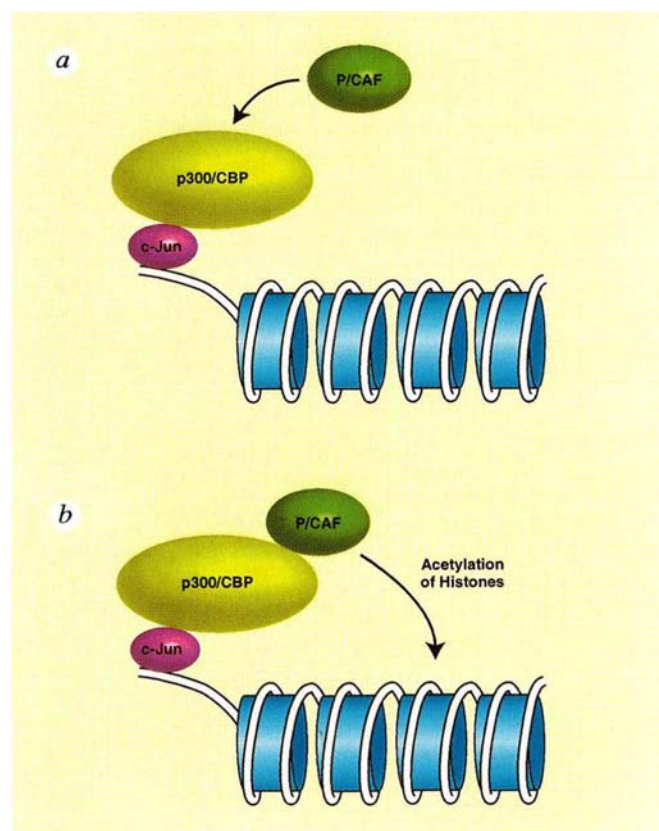


FIG. 6 Model of molecular functions of P/CAF and E1A. p300/CBP could be recruited to promoter elements by interaction with various sequence-specific activators, including c-Jun⁹, Fos¹¹, CREB^{3,4,7}, YY1 (ref. 10), c-Myb¹² and nuclear receptors¹³. P/CAF could also be recruited onto the specific promoters through interaction with p300/CBP (a), then acetylate histones in a promoter-specific manner (b). Targeted histone acetylation could contribute to promoter activation by changing or disrupting the repressive chromatin structure. The E1A oncoprotein perturbs normal cellular regulation between P/CAF and p300/CBP by competing for the binding sites in p300/CBP.

(E1AAN) has very weak activity (Fig. 4e). Comparable expression between wild-type and mutant E1A in transfected cells was revealed by immunoblotting with anti-E1A (data not shown). Intriguingly, when P/CAF was cotransfected with E1A, the mitogenic activity of E1A was significantly counteracted by P/CAF (compare Fig. 4c–f). These results support the view that P/CAF and E1A mediate antagonistic effects on cell-cycle progression.

In the course of assessing P/CAF activity, we also found that p300 is able to inhibit cell-cycle progression under the same assay conditions (V.V.O. and B.H.H., unpublished observation). These findings suggest that P/CAF and p300, perhaps by forming a complex, act together to suppress cell-cycle progression.

Histone acetylase activity in P/CAF

Acetylation of the N-terminal histone tails is considered to be crucial for the accessibility of transcription factors to nucleosomal templates^{26,27}. As yGCN5 has recently been identified as a histone acetylase²⁸, we tested for intrinsic histone acetyltransferase activity in P/CAF and hGCN5. As substrates, we used the core histones (histones H2A, H2B, H3 and H4) and nucleosome core particles (146 base pairs of DNA wrapped around an octamer of core histones). We found that P/CAF and hGCN5 acetylated the core histones with almost the same efficiency (Fig. 5, lanes 1 and 2). Both factors acetylated histones H3 and H4, but H3 preferentially. Very weak or no acetylation by hGCN5 was detected in nucleosome core particles (lane 4), but for P/CAF this was significant (lane 5).

Although all core histones are acetylated in the nucleus, P/CAF and hGCN5 did not acetylate histones H2A and H2B *in vitro* (Fig. 5). P/CAF and hGCN5 may form a heteromeric complex with ADA2 and ADA3 (refs 19–22), and these factors may modulate the substrate specificity of the catalytic subunit. Alternatively, another histone acetylase may acetylate histones H2A and H2B.

Implications

Many oncogene products induce tumour formation by titrating out tumour suppressors. The transforming activity of adenovirus E1A involves two distinct sites in the protein which independently target RB-related proteins and p300/CBP^{1,2}. Interactions of E1A with RB-related proteins and p300/CBP influence functionally distinct growth-regulatory pathways¹. In the former interaction, E1A targets the RB pocket, which also binds to various cellular

factors^{1,2}, including E2F, ATF, and BRG1/hbrm. Thus, E1A probably inactivates the normal growth-suppressing function of RB by competing directly with various RB-associated factors. Although the p300/CBP and RB pathways are targeted by different E1A domains, their mechanisms could be similar. Like the RB pathway, both E1A and P/CAF bind to the same or very closely spaced sites on p300/CBP. Moreover, E1A competitively disrupts complex formation between P/CAF and p300/CBP.

Direct participation of P/CAF is likely to involve its intrinsic histone acetyltransferase activity. Although the exact molecular mechanisms by which acetylation of core histones contributes to transcription are not clear, acetylation of histones is considered to be important in transcriptional regulation^{26,27}. The positively charged N-terminal tails of core histones are believed to affect nucleosome structure by interacting with DNA at or near the nucleosome–spacer junction. Acetylation of the histone tails presumably destabilizes the nucleosome and facilitates access by regulatory factors. Likewise, there is a general correlation between the level of acetylation and transcriptional activity of nucleosomal domains. Our findings fit a scheme of targeted histone acetylation like that shown in Fig. 6: P/CAF could be recruited onto the specific promoters through selective interaction, then acetylate histones in a promoter-specific manner. The SWI/SNF complex, a global gene regulator whose functions include remodelling chromatin, functionally interacts with yGCN5, yADA2, and yADA3 (K. Pollard and C. L. Peterson, personal communication). Targeted histone acetylation might contribute to promoter activation by disrupting the repressed chromatin structure, perhaps in conjunction with the SWI/SNF complex.

p300/CBP binds to various sequence-specific factors that are involved in cell growth and/or differentiation, including CREB^{3,4}, c-Jun⁹, Fos¹¹, c-Myb¹² and nuclear receptors¹³. P/CAF could stimulate the activation function of these factors through promoter-specific histone acetylation. Our present work indicates that E1A perturbs normal cellular regulation by disrupting the connection between p300/CBP and its associated histone acetyltransferase.

Note added in proof: Like hGCN5, yGCN5 also does not acetylate nucleosome core particles (L. E. Brownell and C. D. Allis, personal communication). □

Received 28 February; accepted 12 June 1996.

- Moran, E. *Curr. Opin. Genet. Dev.* **3**, 63–70 (1993).
- Dyson, N. & Harlow, E. *Cancer Surv.* **12**, 161–195 (1992).
- Chrivia, J. C. *et al. Nature* **365**, 855–859 (1993).
- Kwok, R. P. *et al. Nature* **370**, 223–226 (1994).
- Eckner, R. *et al. Genes Dev.* **8**, 869–884 (1994).
- Arany, Z., Sellers, W. R., Livingstone, D. M. & Eckner, R. *Cell* **77**, 799–800 (1994).
- Arany, Z., Newsome, D., Oldread, E., Livingston, D. M. & Eckner, R. *Nature* **374**, 81–84 (1995).
- Lundblad, J. R., Kwok, R. P., Lurance, M. E., Harter, M. L. & Goodman, R. H. *Nature* **374**, 85–88 (1995).
- Anias, J. *et al. Nature* **370**, 226–229 (1994).
- Lee, J. S. *et al. Genes Dev.* **9**, 1188–1198 (1995).
- Banister, A. J. & Kouzarides, T. *EMBO J.* **14**, 4758–4762 (1995).
- Dai, P. *et al. Genes Dev.* **10**, 528–540 (1996).
- Kamei, Y. *et al. Cell* **85**, 403–414 (1996).
- Rochette-Egly, C., Fromental, C. & Chambon, P. *Genes Dev.* **5**, 1200–1211 (1990).
- Struhl, K. *Cell* **50**, 841–846 (1987).
- Engelberg, D., Klein, C., Martinetto, H., Struhl, K. & Karin, M. *Cell* **77**, 381–390 (1994).
- Georgakopoulos, T. & Thireos, G. *EMBO J.* **11**, 4145–4152 (1992).
- Berger, S. L. *et al. Cell* **70**, 251–265 (1992).
- Horiuchi, J., Silverman, N., Marcus, G. A. & Guarente, L. *Molec. cell. Biol.* **15**, 1203–1209 (1995).
- Berger, S. L., Cress, W. D., Cress, A., Triezenberg, S. J. & Guarente, L. *Cell* **61**, 1199–1208 (1990).
- Marcus, G. A., Silverman, N., Berger, S. L., Horiuchi, J. & Guarente, L. *EMBO J.* **13**, 4807–4815 (1994).
- Georgakopoulos, T., Gounalaki, N. & Thireos, G. *Molec. gen. Genet.* **246**, 723–728 (1995).
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. *J. molec. Biol.* **215**, 403–410 (1990).

- Adams, M. D., Soares, M. B., Kerlavage, A. R., Fields, C. & Venter, J. C. *Nature Genet.* **4**, 373–380 (1993).
- Candau, R. *et al. Molec. cell. Biol.* **16**, 593–602 (1996).
- Turner, B. M. *Cell* **75**, 5–8 (1993).
- Wolffe, A. P. *Trends biochem. Sci.* **19**, 240–244 (1994).
- Bownell, J. E. *et al. Cell* **84**, 843–851 (1996).
- Haynes, S. R. *et al. Nucleic Acids Res.* **20**, 2063 (1992).
- Hopp, T. P. *et al. Biotechnology* **6**, 1205–1210 (1988).
- Harlow, E. D. & Lane, D. *Antibodies: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, New York, 1988).
- Giordano, T., Howard, T. H., Coleman, J., Sakamoto, K. & Howard, B. H. *Expl. Cell Res.* **192**, 193–197 (1991).
- Ogryzko, V. V., Hirai, T. H., Shih, C. E. & Howard, B. H. *J. Virol.* **68**, 3724–3732 (1994).
- Minucci, S. *et al. Molec. cell. Biol.* **14**, 360–372 (1994).
- Hayes, J. J., Pruss, D. & Wolffe, A. P. *Proc. natn. Acad. Sci. U.S.A.* **91**, 7817–7821 (1994).
- Brownell, J. E. & Allis, C. D. *Proc. natn. Acad. Sci. U.S.A.* **92**, 6364–6368 (1995).

ACKNOWLEDGEMENTS. We thank A. S. Levine for his support and encouragement; D. Allis for critical unpublished information regarding the cloning of histone acetylase; A. P. Wolffe for histones; R. H. Goodman for CBP plasmids; D. M. Livingston and R. Eckner for p300 plasmids; J. M. Sikela for the NIB2000-5R plasmid and sequence information; K. Ozato for the pCX plasmid; J. Schiltz for advice on histone acetylase assay; T. Howard for advice on transfection; M. L. Avantaggiati for advice on immunoprecipitation; M. J. Swanson, A. G. Hinnebusch, M. Brenner, S. Dimitrov, R. Maraia and S. Minucci for critical comments; D. Pruss for the graphic of the nucleosome; and M.-L. Lanigan for editing the manuscript. This study was supported in part by the Ono Pharmaceutical Co. Ltd. (Y.N.).

CORRESPONDENCE and requests for materials should be addressed to Y.N. (e-mail: yoshi@helix.nih.gov). Sequences have been deposited in the GenBank data base: accession numbers for P/CAF and hGCN5 are U57317 and U57316, respectively.

Crystal structure of the p27^{Kip1} cyclin-dependent-kinase inhibitor bound to the cyclin A–Cdk2 complex

Alicia A. Russo, Philip D. Jeffrey, Andrea K. Patten, Joan Massagué* & Nikola P. Pavletich

Cellular Biochemistry and Biophysics Program, * Cell Biology Program and the Howard Hughes Medical Institute, Memorial Sloan-Kettering Cancer Center, New York, New York 10021, USA

The crystal structure of the human p27^{Kip1} kinase inhibitory domain bound to the phosphorylated cyclin A–cyclin-dependent kinase 2 (Cdk2) complex has been determined at 2.3 Å. p27^{Kip1} binds the complex as an extended structure interacting with both cyclin A and Cdk2. On cyclin A, it binds in a groove formed by conserved cyclin box residues. On Cdk2, it binds and rearranges the amino-terminal lobe and also inserts into the catalytic cleft, mimicking ATP.

COMPLEXES of cyclins with cyclin-dependent kinases (CDKs) play a central role in the control of the eukaryotic cell cycle¹. The discovery of proteins that bind to and inhibit the catalytic activity of cyclin–CDK complexes has identified kinase inhibition as an intrinsic component of cell-cycle control (reviewed in refs 1–3). These inhibitors (CKIs) induce cell-cycle arrest in response to anti-proliferative signals, including contact inhibition and serum deprivation⁴, TGF- β (ref. 5), myogenic⁶, myeloid⁷ and neuronal differentiation⁸, and DNA-damage checkpoints⁹. The inhibitors, which are present in proliferating cells as well, may also help to coordinate cell-cycle progression by their redistribution between different cyclin–CDK complexes^{10,11}.

The inhibitors that have been identified so far can be grouped, according to sequence and functional similarities, into two families. The Kip/Cip family of inhibitors, which include p27^{Kip1} (refs 4, 12), p21^{Cip1, WAF-1} (refs 9, 13–15) and p57^{Kip2} (refs 16, 17), bind to and inhibit cyclin–CDK complexes with broad preference for the G1 and S phase kinase complexes over the mitotic ones. The Kip/Cip inhibitors can bind isolated cyclin and CDK subunits independently, but they have a higher affinity for the cyclin–CDK complexes^{12,18–20}. The INK4 family members, which include p15, p16, p18 and p19, are specific for Cdk4, and its close isoform Cdk6, and can bind to either the isolated CDK subunit or its complex with cyclin D³.

Members of the Kip/Cip family contain a 65-amino-acid region with homology (38–44% identity) at their N-terminal portions, which is necessary and sufficient to bind to and inhibit cyclin–CDK complexes^{4,16,18,21}. Their carboxy-terminal portions are variable in length and divergent in sequence and function³.

Here we report the crystal structure of the 69-amino-acid N-terminal inhibitory domain of p27^{Kip1} bound to the phosphorylated cyclin A–Cdk2 complex, determined at 2.3 Å resolution (Table 1 and Fig. 1). The structure reveals that p27 uses a three-stage approach to bind and inhibit the cyclin A–Cdk2 complex. It binds a peptide-binding groove on the conserved cyclin box of cyclin A, it binds the N-terminal lobe of Cdk2 and it also inserts deep inside the catalytic cleft, mimicking ATP. A comparison of Cdk2 in this complex with the structure of Cdk2 in the binary cyclin A–Cdk2 complex^{22,23} reveals that p27 binding causes large conformational changes in and around the catalytic cleft of Cdk2.

Overall structure of the complex

In the complex (Fig. 2a), the p27 inhibitory domain has a non-globular, extended structure that consists sequentially of a rigid coil, an amphipathic α -helix, an amphipathic β -hairpin, a β -strand and a 3_{10} helix (Fig. 2b). p27 does not have a hydrophobic core of

its own, and its secondary structure elements do not interact significantly with each other.

The extended structure of p27 allows it to cover and interact with a large surface area of the cyclin A–Cdk2 complex. The surface area buried when p27 binds to the cyclin A–Cdk2 complex is 5,752 Å² (40% on cyclin A, 60% on Cdk2), which is significantly larger than other protein–protein complexes, including the cyclin A–Cdk2 complex (3,550 Å²; ref. 22). The majority of the p27 hydrophobic amino acids pack against the cyclin A–Cdk2 complex (Fig. 2b) instead of participating in the folding of p27. The extended p27 structure can thus be thought of as refolding on the cyclin A–Cdk2 complex. This hypothesis is consistent with the large buried surface area per amino acid, which is comparable to that expected from protein folding²⁴.

The structure of the cyclin A subunit consists of two five-helix structural repeats, the first of which corresponds to the conserved cyclin box²². It is only the cyclin-box repeat that is involved in p27 binding (Fig. 2a), interacting with the rigid coil and α -helix of p27. The Cdk2 structure consists of an N-terminal lobe that is rich in β -sheet, and a larger, C-terminal lobe that is mostly α -helical, with the catalytic cleft in between the two lobes²². In the complex, the N-terminal lobe and catalytic cleft of Cdk2 bind the β -hairpin, β -strand, and 3_{10} helix of p27 (Fig. 2a).

The rigid coil of p27, which is 10-amino-acids long (Fig. 2b) and includes the sequence Leu-Phe-Gly conserved among all Kip/Cip family members (henceforth, the 'LFG' motif), binds a shallow groove on cyclin A in a partially extended conformation. This peptide-binding groove is formed by the $\alpha 1$, $\alpha 3$ and $\alpha 4$ helices of the cyclin-box repeat and is lined with amino acids highly conserved among members of the cyclin family (Fig. 2a; ref. 22). This may explain, in part, the ability of p27 to bind many of the cyclin–CDK complexes. The peptide-binding groove of cyclin A revealed by the structure is likely to be the docking site of tight binding substrates, such as p107, which also contain the LFG sequence motif²⁵. Immediately after the LFG coil, the hydrophobic face of the amphipathic p27 helix packs against a surface formed by the $\alpha 4$ and $\alpha 5$ helices of the cyclin box. The interactions made by the helix are not as extensive as those made by the LFG coil, and we presume that they contribute less to the binding. The helix axis is directed towards Cdk2, and as the helix leaves the cyclin A area and encounters a gap on the surface of cyclin A–Cdk2 it becomes kinked and partially disordered (residues 49–53).

As the helix reaches Cdk2, it becomes well ordered again. The helix is followed by an amphipathic β -hairpin (Fig. 2a, b) whose hydrophobic face packs against the N-terminal β -sheet of Cdk2 (strands 3, 4 and 5), in a β -sandwich arrangement. Immediately

after the β -hairpin, a p27 β -strand displaces one from Cdk2 and is incorporated into the Cdk2 β -sheet (Fig. 2). The p27 3_{10} helix that follows, inserts into the catalytic cleft of Cdk2 (Fig. 2). The 3_{10} helix contains the Phe-Tyr pair of amino acids which become buried deep inside the catalytic cleft, and which mimic the purine base of the ATP in their van der Waals and hydrogen bond contacts to Cdk2.

Binding of p27 to Cdk2 is accompanied by extensive structural

changes in the N-terminal lobe and catalytic cleft of Cdk2. The Cdk2 β -sheet flattens and presents hydrophobic side chains to p27. In the binary complex, these side chains interact with each other instead, owing to the large curvature of the β -sheet. A strand of the Cdk2 β -sheet is replaced with one from p27, and this moves another Cdk2 strand by as much as 8.5 Å. These changes, together with the bulky 3_{10} helix binding, widen the catalytic cleft.

Inhibition of kinase activity by p27 thus has several components.

TABLE 1 Statistics from the crystallographic analysis

Data set	Native	HgCl ₂	Thimerosal	Uranyl acetate
Resolution (Å)	2.3	2.8	3.0	3.2
Observations	162339	124684	65378	21031
Unique reflections	34683	17569	15964	10528
Data coverage (%)	96.1	86.1	96.0	77.0
R_{sym} (%)	5.6	7.2	7.3	9.1
MIR analysis (20.0–3.2 Å):				
Mean isomorphous difference		0.18	0.29	0.17
Phasing power		1.4	0.9	0.7
Refinement:				
Resolution range (Å)	7.0–2.3			
R-factor	19.2			
Reflections with $ F > 2\sigma$	31351			
Number of atoms	5181			
Number of waters	197			
R.m.s.* bond lengths (Å)	0.012			
R.m.s.* bond angles (deg)	1.45			
R.m.s.* B-factors (Å ²)	3.26			

Preparation of ternary complex. The structural organization of the p27 protein⁴ was characterized using subtilisin digestion which yielded an N-terminal fragment (residues 28–96) that contained CDK-inhibitory activity essentially identical to that of full-length p27, and a C-terminal fragment (residues 123–198). Recombinant human p27 inhibitory domain, corresponding to residues 22–106, was overexpressed in *Escherichia coli* using the pET3a vector (Novagen) according to standard procedures. The soluble fraction of the *E. coli* lysate was clarified over a Q-Sepharose column, the flow-through of which was applied to a heparin sulphate column at pH 6.0. The p27 fragment was eluted with a NaCl gradient, concentrated by ultrafiltration, and further purified by gel-filtration chromatography. Differential scanning calorimetry (data not shown) showed that the isolated p27 domain was not folded, suggesting that it may need the C-terminal domain to remain folded. Alternatively, the isolated inhibitory domain may be in partly folded states that are resistant to digestion, but which do not represent a single state that can be detected calorimetrically. To prepare the p27–cyclin A–Cdk2 complex, a twofold molar excess of the recombinant inhibitory domain was mixed with the phosphorylated cyclin A–Cdk2 complex, prepared as described²³, at a total protein concentration of 13 mg ml^{−1} (determined by ultraviolet absorption at 280 nm, calibrated using standard extinction coefficients for Phe, Tyr and Trp amino acids). The mixture was fractionated by gel-filtration chromatography, yielding a peak at a retention volume corresponding to a relative molecular mass of 70K containing the ternary complex at an ~1:1:1 molar ratio, and a second peak containing a onefold molar excess of free p27. The ternary complex peak has H1 histone kinase activity about 1% that of the phosphorylated binary complex²³ when assayed at a concentration of 10 nM, which is about an order of magnitude above the K_i of p27 for the related cyclin E–Cdk2 complex²⁰. The residual activity probably results from the destabilization of the complex owing to its low concentration, because the residual activity can be made arbitrarily small by performing the kinase assay at concentrations well above the K_i (data not shown). The ternary complex peak was concentrated by ultrafiltration to 30 mg ml^{−1} in a buffer of 40 mM HEPES, 200 mM NaCl and 5 mM dithiothreitol, pH 7.0.

Crystallization and data collection. Crystals were grown at 4 °C by the hanging-drop vapour-diffusion method, by mixing the complex with an equal volume of the well buffer containing 35% saturated ammonium sulphate, 40 mM HEPES, 5 mM dithiothreitol, pH 7.0. The crystals form in space group $P2_12_12_1$ with $a = 73.8$, $b = 78.3$, $c = 137.2$ Å and contain one complex in the asymmetric unit. Diffraction data were collected at −170 °C (ref. 22) using an R-Axis IIc imaging plate detector mounted on a Rigaku 200HB generator. Heavy-atom soaks were performed in crystallization buffer lacking dithiothreitol and containing one of the following heavy-atom solutions: 0.6 mM HgCl₂ for 3 h, 0.8 mM thimerosal for 4 h, 2.5 mM uranyl acetate for 2 d. Molecular replacement and MIR analysis. The structure was determined by a combination of molecular replacement (MR) and multiple isomorphous replacement (MIR) methods. The positions of cyclin A and Cdk2 were determined by MR with the program AMORE³², using the structure of the unphosphorylated binary complex²². The correlation coefficient was 0.59 for 10–4 Å data. Maps calculated from the MR solution had clear electron density for the LFG motif (Fig. 1) and the 3_{10} helix of p27. To obtain phases with reduced model bias, an MIR analysis was undertaken using heavy-atom sites identified with difference Patterson and model-phased difference Fourier methods. The MIR analysis with the CCP4 program suite³³ included anomalous scattering from the HgCl₂ and thimerosal derivatives and had a mean figure of merit of 0.56 for 20.0–3.2 Å data (nonisomorphism limited the use of derivative data to low resolution). Maps calculated from phases generated by combination of MIR and model phases showed continuous electron density for p27, with detailed features of p27 electron density becoming evident in α_{calc} maps upon further refinement. The structure was refined by simulated annealing with the program X-PLOR³⁴, using the free-R as a monitor of model quality (the final free-R was 26.9 for 7.0–2.3 Å data), followed by least-squares refinement with the program TNT³⁵. The final model consists of residues 13–298 of Cdk2, residues 175–432 of cyclin A, residues 25–93 of p27, one phosphothreonine, one sulphate ion and 197 water molecules. A comparison with the structure of the phosphorylated cyclin A–Cdk2 complex²³ shows that cyclin A does not undergo any significant conformational changes. $R_{\text{sym}} = \sum_i \sum_j |I_{h,i} - I_{h,j}| / \sum_i \sum_j I_{h,i}$ for the intensity (I) of i observations of reflection h ; mean isomorphous difference = $\sum |F_{\text{PH}} - F_{\text{P}}| / \sum F_{\text{PH}}$, where F_{PH} and F_{P} are the derivative and native structure factors, respectively; phasing power = $[(F_{\text{H(calc)}}^2 / (F_{\text{PH(obs)}} - F_{\text{PH(calc)}})^2)^{1/2}]$; R factor = $\sum |F_{\text{obs}} - F_{\text{calc}}| / \sum |F_{\text{obs}}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively.

* Root-mean-square deviations from ideal geometry and root-mean-square variation in the B-factors of bonded atoms.

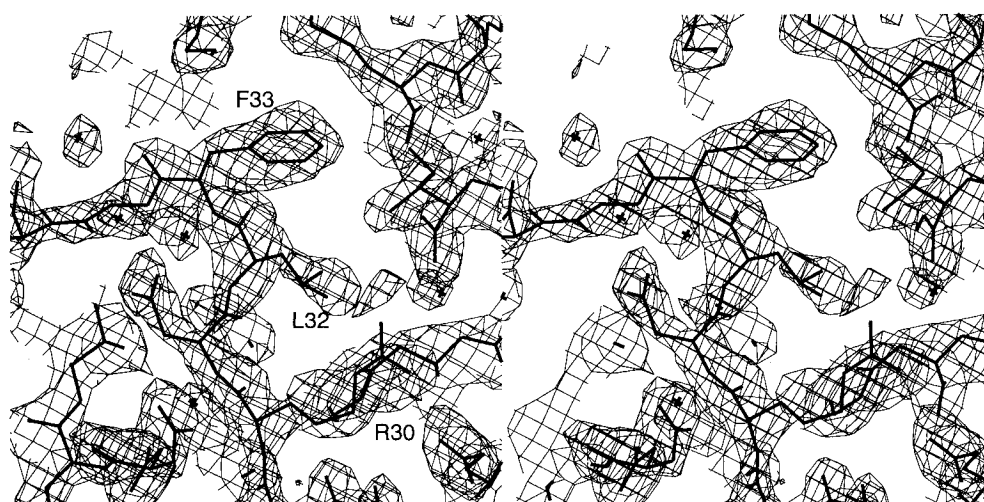


FIG. 1 Initial electron density of the LFG motif at the p27–cyclin A interface contoured at 1.0σ . The $(2|F_{\text{obs}}| - |F_{\text{calc}}|)$ Fourier synthesis was calculated at $2.3\text{ \AA}$ resolution using phases derived from the cyclin A–Cdk2 molecular replacement model before any p27 was built into the model. Arg 30, Leu 32 and Phe 33 of p27 are labelled.

First, by rearranging the N-terminal β -sheet and the catalytic cleft, p27 would destabilize any ATP binding. Second, p27 fills up the catalytic cleft mimicking ATP, and eliminates any possibility of ATP binding. Third, p27 binds a peptide-binding groove of cyclin A, which is likely to be the docking site of a number of tight-binding substrates, also reducing the possibility that substrates may compete away the changes in the catalytic cleft of Cdk2.

p27–cyclin A interactions

The p27–cyclin A interactions are likely to be as important for binding as the p27–Cdk2 interactions because they account for 40% of the surface area buried, they include 35% of the short-distance hydrogen bonds, and they involve residues highly conserved among both the cyclin and p27 families. The most striking p27–cyclin A interactions involve the LFG coil binding a shallow

groove of cyclin A consisting of the $\alpha 1$, $\alpha 3$ and $\alpha 4$ helices of the cyclin-box repeat (Fig. 3). The coil has a rigid conformation and low temperature factors in the crystals, due, in part, to intramolecular hydrogen bonds, in addition to cyclin A contacts. At its N terminus, the coil starts with a reverse turn (residues 26 to 29), followed by a short extended region (residues 29 to 31), and ending with another turn (residues 31 to 34) stabilized by an intramolecular side-chain–backbone hydrogen bond (Fig. 3a). The first turn and the extended region bind cyclin A primarily through hydrogen bonds, making 1 backbone–backbone, 3 side-chain–backbone and 3 side-chain–side-chain hydrogen bonds (see Fig. 3a and legend for a list). These interactions are augmented by van der Waals contacts made by Ala 28 to residues

The second turn contains the Leu 32 and Phe 33 residues of the

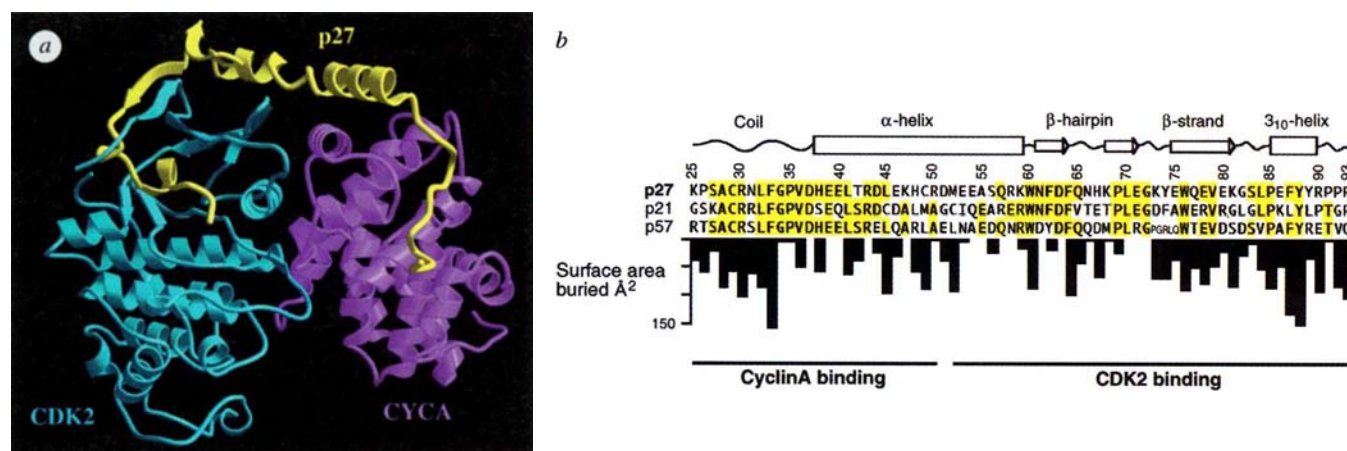


FIG. 2 a, p27 binds as an extended structure interacting with both cyclin A and Cdk2. Schematic drawing shows p27 in yellow, cyclin A in purple, and Cdk2 in cyan. The figure was prepared with the programs MOLSCRIPT³⁶ and RASTER3D³⁷. b, Sequence of the p27 inhibitory domain showing its secondary structure elements and the homology with the corresponding domains of the p21 and p57 Kip/Cip family members (p21 starts at residue

14, and p57 at residue 26). Residues identical in two or more family members are highlighted in yellow. The bar graph shows the buried surface area per p27 amino acid upon complex formation. The structure of unbound p27 is not known, so for the surface accessibility calculations it was assumed to have the same structure as in the complex.

LFG motif, which bind deeper inside the cyclin A groove (Fig. 3b). This area of the cyclin A groove is lined with hydrophobic amino acids highly conserved among cyclins A, B, D1 and E (Met 210, Ile 213, Leu 214, Trp 217 of the $\alpha 1$ helix, and Leu 253, Arg 250 and Gln 254 of the $\alpha 3$ helix), and these amino acids make multiple van der Waals contacts to the Leu-Phe pair of p27 (Fig. 3b). The contacts to the $\alpha 1$ helix are particularly significant because the cyclin A residues involved are among the most highly conserved in the cyclin box (in bold: **MRAILVDW**).

The glycine of the LFG motif (Gly34) has backbone ϕ - ψ angles that other amino acids cannot adopt, and in conjunction with the following proline (Pro35), plays an important role in allowing the polypeptide chain to exit the cyclin A groove. After its exit from the groove, an amphipathic helix (His38, Leu41, Leu45 and Cys49 form the hydrophobic face of the helix) guides the chain towards Cdk2, in the process making a few additional van der Waals contacts with the $\alpha 5$ helix on the surface of cyclin A.

p27-Cdk2 interactions

These interactions involve the N-terminal lobe of Cdk2, which in the binary cyclin A-Cdk2 complex²² contains a five-stranded β -sheet, as well as the catalytic cleft of Cdk2 occurring at the interface of the N- and C-terminal lobes. p27 uses three consecutive secondary structure elements, its β -hairpin, β -strand and 3_{10} helix, to clamp around the β -sheet of Cdk2 (Fig. 2). These three regions of p27 are likely to bind in a highly interdependent and cooperative fashion, because they rely on structural changes in CDK2 which they together induce.

First, the β -hairpin (residues 61–71), which is amphipathic, forms a sandwich with the N-terminal β -sheet of Cdk2, burying a large number of hydrophobic and aromatic amino acids both from p27 and Cdk2 (Fig. 4a, b). In this region of p27, the interacting amino acids include Trp60, Phe62, Phe64, Pro69, Tyr74 and Trp76, and they are highly conserved in the Kip/Cip family (Fig. 2b). On Cdk2, the interacting amino acids include Tyr19, Ala21, Val30, Leu32, Leu67, Ile70, Tyr77 and Val79, and they

are generally conserved as hydrophobic amino acids in the Cdk family. In the binary cyclin A-Cdk2 complex, these hydrophobic Cdk2 amino acids instead pack against each other, their interactions being facilitated by the β -sheet being highly curved and partly folded upon itself (Fig. 4c). On p27 binding, the β -sheet flattens and exposes the hydrophobic amino acids to p27. In addition to the van der Waals contacts, this region of the interface also includes two backbone-side-chain hydrogen bonds, and two side-chain-side-chain hydrogen bonds (Fig. 4a legend).

The second portion of the p27-Cdk2 interface involves the p27 β -strand (residues 75–81) being incorporated into the Cdk2 β -sheet through six backbone-backbone β -sheet hydrogen bonds (with residues 16–22 of Cdk2 in an antiparallel arrangement). To accommodate the p27 strand, the first β -strand of Cdk2 (residues 1–13) is displaced and becomes disordered (Fig. 4a, b). This results in a p27-Cdk2 hybrid β -sheet that has four strands from Cdk2 and one strand from p27.

The third portion of the p27-Cdk2 interface involves the 3_{10} helix of p27 (residues 85–90) that binds deep inside the catalytic cleft and occupies most of the available space in the cleft (Fig. 5a). The helix axis is parallel with the long dimension of the cleft, and the helix traces the position that the ATP occupies in the binary complex^{22,23}. In this region, the most significant contacts are made by Tyr88, which is conserved in the Kip/Cip family (Fig. 2b). Its aromatic ring makes van der Waals contacts to Phe80, Phe82 and Leu134 of Cdk2, and its hydroxyl group makes two hydrogen bonds to the backbone carbonyl of Glu81 and the backbone amide of Leu83 (Fig. 5a). The position of the tyrosine side chain in the catalytic cleft, and the Cdk2 contacts it makes, mimic those that are made by the purine base of ATP in the binary complex (Fig. 5b; ref. 22). In particular, the tyrosine hydroxyl group hydrogen bonds are analogous to the hydrogen bonds made by the N1 and N6 groups of the adenine base (Fig. 5b).

Other packing interactions in this region involve Phe87, which is also positioned deep inside the catalytic cleft, making van der Waals contacts to Phe80, Leu134 and Ala144 of Cdk2 (Fig. 5a).

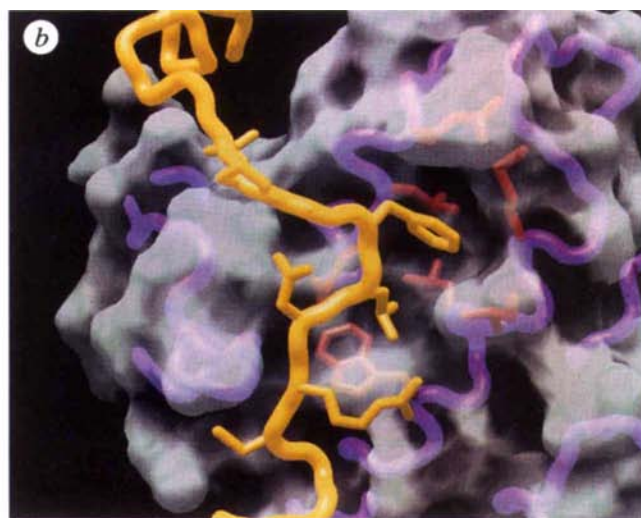
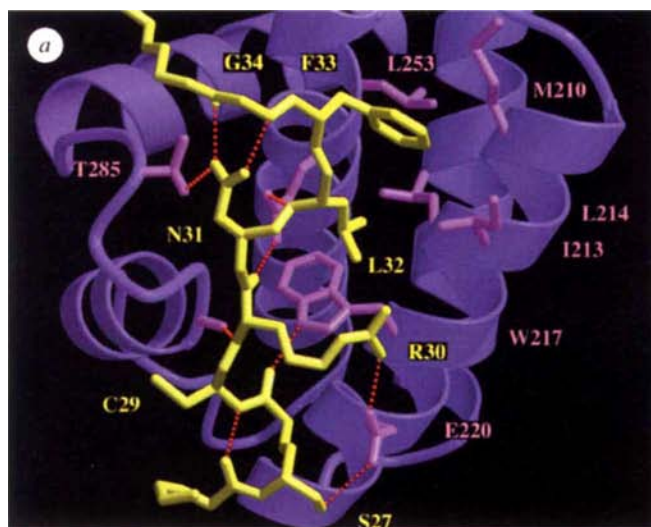


FIG. 3 a, The rigid coil of p27 binds a shallow cyclinA groove. View highlights the multiple hydrogen-bond (red dotted lines) and van der Waals contacts at the interface. The orientation and colouring are similar to those of Fig. 2a, except that side chains are coloured in a brighter tone than the backbone representations. On p27, side chains for residues 27–34 are shown, except for Ala28 whose side chain points into the plane of the figure and is obstructed in this view; on cyclinA, only the interacting side chains are shown. Following is a complete list of intermolecular hydrogen bonds in this region: Ser27 O_γ of p27 hydrogen bonds with Glu220 O_{ε2} of

cyclinA, Ala28 carbonyl with Trp217 N_{ε1}, Arg30 amide with Ile281 carbonyl, Arg30 N_{γ1} with Glu220 O_{ε1}, Arg30 carbonyl with Gln254 N_{ε2}, Asn31 N_{δ2} with Thr285 O_{γ1}, and Leu32 amide with Gln254 O_{ε1}. b, Surface representation of cyclinA, highlighting the concave parts of the surface (in grey) that form the binding sites of the hydrophobic p27 amino acids in an orientation similar to that in a. Side chains shown in a are also shown here. The cyclinA surface was constructed with the program GRASP³⁸.

In addition to the van der Waals contacts, the 3_{10} helix also forms hydrogen bonds with conserved catalytic-site amino acids of Cdk2. The backbone carbonyl groups of Phe 87 and Arg 90 of p27 make hydrogen bonds to the conserved Lys 33 side chain of Cdk2. The positions of the Phe 87 and Arg 90 carbonyl groups are similar to those of the ATP phosphates in the binary complex, and the carbonyl–Lys 33 interactions mimic the ATP phosphate–Lys 33 interactions of the binary complex (Fig. 5b).

Structural changes in Cdk2

The five stranded β -sheet of Cdk2, which forms the roof of the catalytic cleft and plays a critical role in ATP binding²², is altered by p27 binding in several ways. First, in the binary complex, the β -sheet is curved and partially folded upon itself burying a large number of hydrophobic amino acids (Fig. 4c). On p27 binding, the displacement of the first strand provides partial access to the hydrophobic patch by removing the Met 1 and Phe 4 side chains from the packing (Fig. 4a, c). The straight p27 strand that replaces

it moves the N-terminal portion of the second Cdk2 strand by up to 8.5 Å, separating it partially from the rest of the β -sheet (Fig. 4a, c). These changes, along with the overall flattening of the β -sheet, are associated with the exposure of the hydrophobic patch and its repacking with the amphipathic p27 β -hairpin.

Second, in the binary complex, the loop between β -strands 1 and 2 (the glycine loop) folds partially over the ATP-binding portion of the cleft and binds the ATP phosphates^{22,23}. On p27 binding, the glycine loop is eliminated, and the position that the glycine loop would occupy now has the coil immediately after the 3_{10} helix of p27. Furthermore, the partial separation of the second β -strand from the rest of the Cdk2 β -sheet also creates an opening in the roof of the catalytic cleft. These changes, which effectively open up the ATP-binding portion of the catalytic cleft, are accompanied by the binding of the p27 3_{10} helix, which is significantly bulkier than ATP, in the catalytic cleft.

The changes in the β -sheet, and in particular in the ATP-binding glycine loop, suggest that even in the absence of the 3_{10}

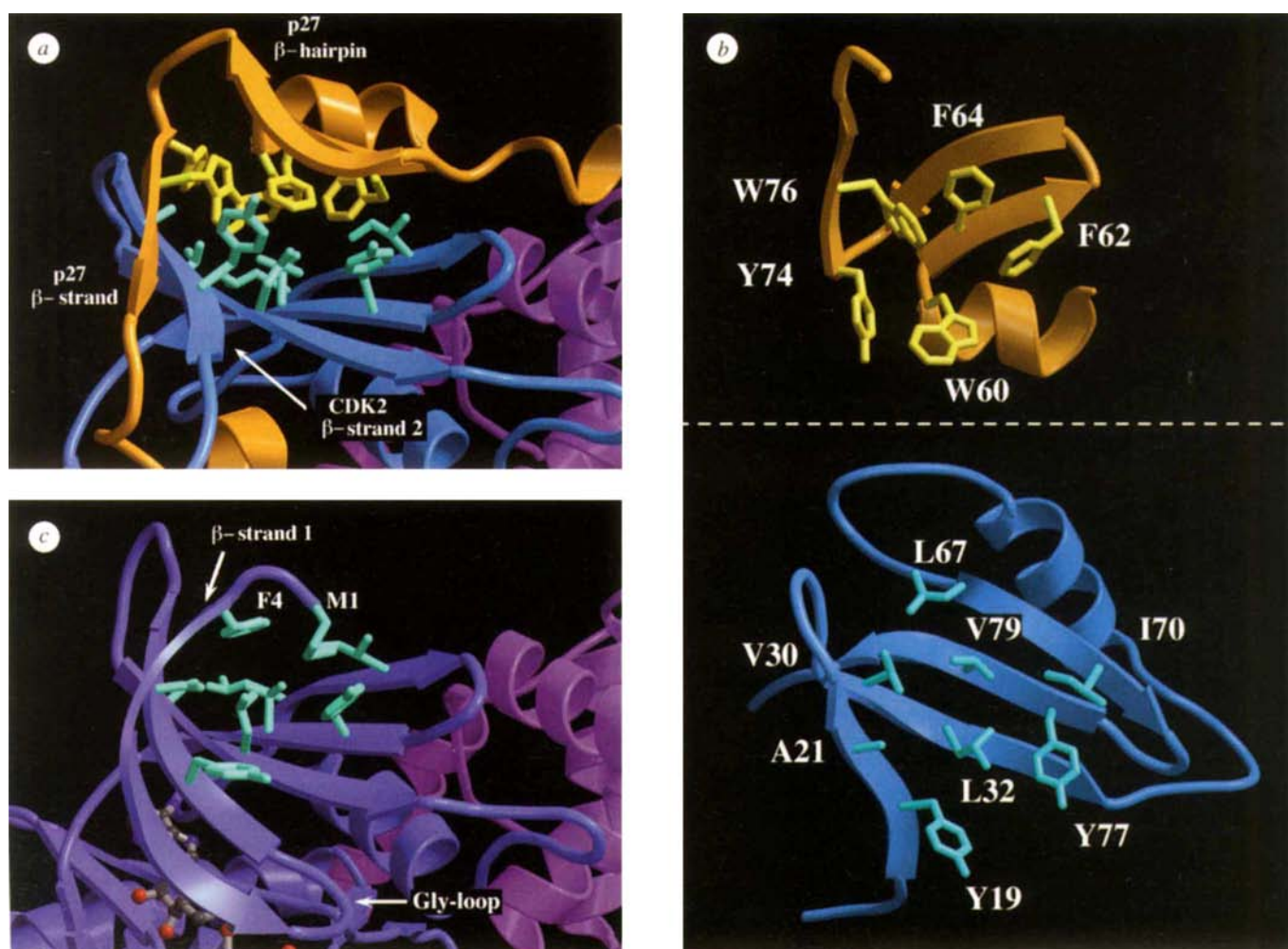


FIG. 4 a, The p27 β -hairpin forms a sandwich with the N-terminal β -sheet of Cdk2, resulting in changes in the structure of the N-terminal lobe of Cdk2. The orientation and colouring are similar to those of Fig. 2a. Additional contacts not mentioned in the text include hydrogen bonds from the Lys 75 side chain of Cdk2 to the backbone carbonyls of Phe 64 and Gln 65 of the β -hairpin tip, and hydrogen bonds from the Trp 60 and His 67 side chains of Cdk2 to the Asp 68 and Tyr 77 side chains of p27, respectively. b, The β -sandwich is opened up about the dotted line to show the hydrophobic amino

acids, which are labelled. c, In the phosphorylated cyclin A–Cdk2 complex^{22,23} the β -sheet is curved and partly folded upon itself, burying the hydrophobic amino acids which become exposed on p27 binding (Met 1, Phe 4, Tyr 19, Ala 21, Val 30, Leu 32, Leu 67, Ile 70, Tyr 77 and Val 79). The first strand of the β -sheet and the glycine-rich loop that folds over the ATP phosphates^{22,23}, are eliminated upon p27 binding. The Met1 and Phe4 residues are labelled; ATP is shown as a ball-and-stick representation.

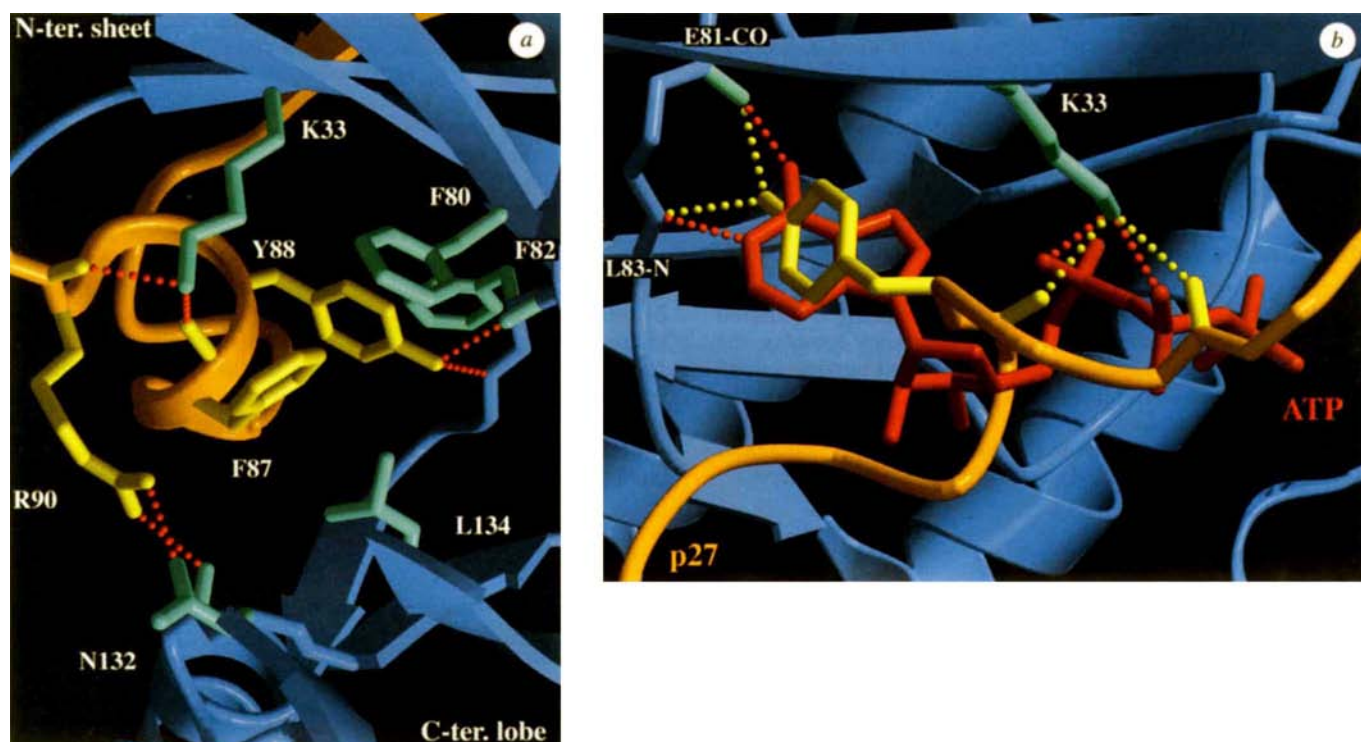


FIG. 5 *a*, The p27 3_{10} helix binds in the widened catalytic cleft of Cdk2 making multiple van der Waals and hydrogen-bond contacts (red dotted lines). View is rotated by roughly 90° about the vertical axis of Fig. 2*a*; colouring is as Fig. 2*a*. To make the contacts easier to see, parts of the N- and C-terminal lobes that would be above the plane of the figure are not shown. Hydrogen-bond contacts not mentioned in the text include ones by the Arg90 side chain of p27 to the Asn132 side chain and Gln131 backbone carbonyl group of Cdk2. *b*, The p27 3_{10} helix mimics the ATP in the binary complex in its position and its Cdk2 contacts. Composite figure

shows the ATP (red) from the phosphorylated binary complex²³ positioned by superimposing the Cdk2 molecules in the binary and tertiary complexes. The hydrogen bonds that ATP makes in the binary complex are indicated as red dotted lines and those made by p27 as yellow dotted lines. View is rotated by roughly 90° about the horizontal axis of Fig. 2*a*, so that the N-terminal β -sheet of Cdk2 is above the plane of the figure. The second strand of the β -sheet is not shown because it would obstruct this view of the catalytic cleft.

helix filling up the catalytic cleft, the kinase would have a reduced affinity for ATP. This is consistent with the observation that a truncated form of p27, lacking the 3_{10} helix residues, retains partial inhibitory activity^{4,21}.

Implications for Kip/Cip family function

The crystal structure of the p27–cyclin A–Cdk2 complex reveals that p27 has separate binding sites on the cyclin and CDK subunits and this explains how the Kip/Cip inhibitors can bind the isolated subunits. Binding to the cyclin–CDK complex is significantly tighter^{12,18–20}, consistent with cooperative binding to the two subunits.

LFG–cyclin A interactions. The LFG motif–cyclin interactions are likely to serve as an initial anchor in complex formation, because these interactions do not need structural rearrangements in cyclin to occur, in contrast to the Cdk2 interactions. The importance of the LFG–cyclin interactions is supported by the observation that a 25-amino-acid peptide encompassing the LFG motif of p21 can antagonize p21 binding to, and inhibition of, the cyclin–CDK complex²⁶.

The LFG–cyclin interactions involve conserved residues on both partners. The LFG motif corresponds to the region best conserved in the Kip/Cip family (Fig. 2*b*), and the region of cyclin A that forms the binding groove corresponds to the start of the cyclin box, also the region best conserved among cyclins A, B, D and E (ref. 22). This explains, in part, the specificity of the

Kip/Cip proteins for complexes containing these cyclins. On the other hand, cyclin H (ref. 27) and p35 (ref. 8), which are CDK activators that do not bind the Kip/Cip inhibitors^{28,29}, are only distantly related to the cyclin-box cyclins, and do not have the p27-interacting amino acids conserved in the other cyclins. Cyclin H is a partner of Cdk7 in a CDK-activating kinase²⁷, and p35 is a partner of Cdk5 involved in neurite outgrowth^{8,30}, and their roles are consistent with their need to avoid inhibition by the Kip/Cip proteins.

The identification of a peptide-binding groove on cyclin A raises the possibility that other cyclin-interacting proteins may use peptide–groove interactions in their binding. Studies of the p107 CDK substrate have shown that it contains a sequence similar to the LFG motif that is necessary for its tight binding to the cyclin A–Cdk2 complex²⁵. It is conceivable that variants of the LFG motif, perhaps not immediately recognizable from sequence alignment, may also be used by other proteins to bind the cyclin groove.

p27–Cdk2 interactions. The interactions with the cyclin may serve as an entry point for the Kip/Cip proteins into the cyclin–CDK complex, but the interactions with the CDK and the accompanying structural changes in the catalytic cleft are likely to be key determinants of inhibition. The inhibition constants (K_i) of the Kip/Cip inhibitors for different cyclin–CDK complexes vary considerably²⁰, even for complexes that have consensus residues in their cyclin grooves, and this modulation is likely to reside with the

CDK subunit. For example, the mitotic cyclin B-Cdc2 is only poorly inhibited by the Kip/Cip proteins²⁰, and this is probably due, in part, to a 35-amino-acid insertion at the N terminus of Cdc2. The insertion occurs within a region that corresponds to the first β -strand of Cdk2, and it is likely to interfere with the structural rearrangements in the N-terminal lobe that are necessary for CDK binding and inhibition. Furthermore, we note that the Cdk2 amino acids that make significant contacts to p27 are generally conserved in CDKs 4, 5 and 6, which are also targets of the Kip/Cip proteins, but not in Cdc2 (for example Lys 75 of Cdk2; Fig. 5a legend).

Kip/Cip binding to the cyclin-CDK complex also inhibits the phosphorylation of the CDK subunit by CAK^{4,28,29}, and this is probably a result of steric hindrance. In the crystal structure, which contains phosphorylated Cdk2 (ref. 23), the C terminus of the p27 inhibitory domain is within 8 Å of the phosphothreonine 160 on the T-loop of Cdk2, and a comparison with the unphosphorylated cyclin A-Cdk2 complex²², which has a different T-loop conformation, suggests that p27 would be even closer to the unphosphorylated Thr 160. Full-length p27 contains an additional 105 residues following its inhibitory domain (p21 and p57 contain 82 and 220 additional residues, respectively), and we presume that a significant portion of these amino acids would be in the vicinity of the T-loop as the p27 chain exits the catalytic cleft of Cdk2 and may interfere with the binding of CAK to the T-loop.

Implications of two binding sites. The existence of two distinct binding sites on the cyclin-CDK complex could, in principle, allow the association of two inhibitor molecules with a single cyclin-CDK complex, one interacting with the cyclin, and the other with the CDK. However, this mode of binding will not be cooperative, and the equimolar complex should be greatly pre-

ferred. It has been proposed that there may be cyclin-CDK complexes associated with multiple inhibitor molecules^{20,31}. In the case of cyclin A-Cdk2, we have not detected stable complexes containing multiple p27 molecules, even at high concentrations (up to 0.2 mM) and in the presence of excess p27 (Table 1). We surmise that in the case of cyclin A-Cdk2, two-inhibitor complexes are thermodynamically highly unfavoured over the equimolar complex, although their existence with other Kip/Cip-cyclin-CDK combinations remains a possibility.

The existence of two distinct binding sites could also allow, in principle, weak, unproductive association of a Kip/Cip inhibitor with a cyclin-CDK complex without significant inhibition of the catalytic activity^{20,31}. Such a complex would require a cyclin that has a consensus LFG-motif binding site and a CDK subunit that is not susceptible to inhibitor binding, conditions that might theoretically be met by cyclin-CDK complexes other than cyclin A-Cdk2.

Conclusion

The inhibitory domain of p27 binds the cyclin A-Cdk2 complex as an extended structure interacting with both subunits. Key hydrophobic and hydrogen-bond interactions between a 10-amino-acid region of p27, which contains the conserved LFG sequence motif, and a shallow groove of cyclin A, which consists of conserved cyclin-box residues, are likely to serve as an anchor point in initial binding, facilitating subsequent interactions with Cdk2. p27 binding inhibits the CDK catalytic activity because its interactions with the N-terminal β -sheet of Cdk2 induce conformational changes that change the shape of the catalytic cleft and because p27 inserts into and fills up the catalytic cleft, eliminating any potential for ATP binding. □

Received 16 May; accepted 27 June 1996.

1. Pines, J. *Cancer Biol.* **6**, 63–72 (1995).
2. Peter, M. & Herskowitz, I. *Cell* **79**, 181–184 (1994).
3. Sherr, C. J. & Roberts, J. M. *Genes Dev.* **9**, 1149–1163 (1995).
4. Polyak, K. et al. *Cell* **78**, 59–66 (1994).
5. Reynisdóttir, I., Polyak, K., Iavarone, A. & Massagué, J. *Genes Dev.* **9**, 1831–1845 (1995).
6. Parker, S. B. et al. *Science* **267**, 1024–1027 (1995).
7. Liu, M., Lee, M. H., Cohen, M., Bommakanti, M. & Freedman, L. P. *Genes Dev.* **10**, 142–153 (1996).
8. Lee, M. H. et al. *Proc. natn. Acad. Sci. U.S.A.* **93**, 3259–3263 (1996).
9. El-Deiry, W. S. et al. *Cell* **75**, 817–825 (1993).
10. Soos, T. J. et al. *Cell Growth Differ.* **7**, 135–146 (1996).
11. Poon, R. Y. C., Toyoshima, H. & Hunter, T. *Molec. Biol. Cell* **6**, 1197–1213 (1995).
12. Toyoshima, H. & Hunter, T. *Cell* **78**, 67–74 (1994).
13. Xiong, T. et al. *Nature* **366**, 701–704 (1993).
14. Gu, Y., Turck, C. W. & Morgan, D. O. *Nature* **366**, 707–710 (1993).
15. Harper, J. W., Adami, G. R., Wei, N., Keyomarsi, K. & Elledge, S. J. *Cell* **75**, 805–816 (1993).
16. Lee, M. H., Reynisdóttir, I. & Massagué, J. *Genes Dev.* **9**, 639–649 (1995).
17. Matsuo, S. et al. *Genes Dev.* **9**, 650–662 (1995).
18. Chen, J., Jackson, P. K., Kirschner, M. W. & Dutta, A. *Nature* **374**, 386–388 (1995).
19. Lin, J., Reichner, C., Wu, X. & Levine, A. J. *Molec. Cell Biol.* **16**, 1786–1793 (1996).
20. Harper, J. W. et al. *Molec. Biol. Cell* **6**, 387–400 (1995).
21. Luo, Y., Hurwitz, J. & Massagué, J. *Nature* **375**, 159–161 (1995).
22. Jeffrey, P. D. et al. *Nature* **375**, 159–161 (1995).
23. Russo, A. R., Jeffrey, P. D. & Pavletich, N. P. *Nature struct. Biol.* (in the press).
24. Chothia, C. J. *molec. Biol.* **105**, 1–14 (1976).

25. Zhu, L., Harlow, E. & Dynlacht, D. *Genes Dev.* **9**, 1740–1752 (1995).
26. Chen, I. et al. *Oncogene* **12**, 595–607 (1996).
27. Fisher, R. P. & Morgan, D. O. *Cell* **78**, 713–724 (1994).
28. Kato, J., Matsuo, M., Polyak, K., Massagué, J. & Sherr, C. J. *Cell* **79**, 487–496 (1994).
29. Aprelikova, O., Xiong, Y. & Liu, E. T. J. *J. Biol. Chem.* **270**, 18195–18197 (1995).
30. Nikolic, M., Dudek, H., Kwon, Y. T., Ramos, Y. K. & Tsai, L. H. *Genes Dev.* **10**, 816–825 (1996).
31. Zhang, H., Hannon, G. J. & Beach, D. *Genes Dev.* **8**, 1570–1758 (1994).
32. Navaza, J. *Acta Crystallogr.* **A50**, 157–163 (1994).
33. Collaborative Computational Project No. 4. *Acta Crystallogr.* **D50**, 760–763 (1994).
34. Brunger, A. T. *X-PLOR, a System for Crystallography and NMR Version 3.0 Manual* (Yale Univ. Press, New Haven, CT, 1991).
35. Tonrud, D. E., Ten Eyck, L. F. & Matthews, B. W. *Acta crystallogr.* **A43**, 489 (1987).
36. Kraulis, P. J. *J. appl. Crystallogr.* **24**, 946–950 (1991).
37. Merrit, E. A. & Murphy, M. E. *Acta crystallogr.* **D50**, 869–873 (1994).
38. Nicholls, A., Sharp, K. A. & Honig, B. *Proteins Struct. Funct. Genet.* **11**, 281–296 (1991).

ACKNOWLEDGEMENTS. A.A.R. and P.D.J. contributed equally to this work. We thank S. Geromanos of the Sloan-Kettering Microchemistry Facility for N-terminal sequence and mass spectroscopic analyses; K. Polyak for help with the inhibition assays; D. O. Morgan for providing us with the Cdk7 and cyclin H-expressing baculovirus vectors; and R. Kenny for administrative help. Supported by the NIH, the Pew Charitable Trusts, the Arnold and Mabel Beckman Foundation, the Dewitt Wallace Foundation, and the Samuel and May Rudin Foundation.

CORRESPONDENCE and requests for materials should be addressed to N.P.P. (e-mail: Nikola@xray2.mskcc.org). Coordinates have been deposited with the Brookhaven Protein Data Bank (ID code USU).

Laboratory simulation of cosmic string formation in the early Universe using superfluid ^3He

C. Bäuerle*, Yu. M. Bunkov*, S. N. Fisher*, H. Godfrin* & G. R. Pickett†

* Centre de Recherches sur les Très Basses Températures, Centre National de la Recherche Scientifique, BP 166, 38042 Grenoble Cedex 9, France

† School of Physics and Chemistry, Lancaster University, Lancaster LA1 4YB, UK

TOPOLOGICAL defects in the geometry of space-time (such as cosmic strings) may have played an important role in the evolution of the early Universe, by supplying initial density fluctuations which seeded the clusters of galaxies that we see today¹. The formation of cosmic strings during a symmetry-breaking phase transition shortly after the Big Bang is analogous to vortex creation in liquid helium following a rapid transition into the superfluid state; the underlying physics of this cosmological defect-forming process (known as the Kibble mechanism¹) should therefore be accessible to experimental study. Superfluid vortices have been observed in ^4He following rapid quenching to the superfluid state², lending qualitative support to Kibble's contention that topological defects are generated by such phase transitions. Here we quantify this process by using an exothermic neutron-induced nuclear reaction to heat small volumes of superfluid ^3He above the superfluid transition temperature, and then measuring the deficit in energy released as these regions of normal liquid pass back into the superfluid state. By ascribing this deficit to the formation of a tangle of vortices, we are able to infer the resulting vortex density; we find that this agrees very well with the predictions of Zurek's modification³ of the original Kibble mechanism¹.

In comparison with superfluid ^4He and liquid crystals, superfluid ^3He provides the ideal system for modelling the formation and evolution of topological defects remaining from the cascade of symmetry-breaking phase transitions which the Universe is believed to have undergone shortly after the Big Bang. In superfluid ^4He only gauge symmetry is broken, and in liquid crystals only orbital rotation symmetry is broken. However, owing to the spin and orbital angular momentum properties, ^3He shows a superposition of broken spin rotation, broken orbital rotation and broken gauge symmetries ($\text{SO}^s_3 \times \text{SO}^L_3 \times \text{U}(1)$) which provide a much closer approximation to the superposition of broken rotational and gauge symmetries used to describe the Universe. In fact the analogies between the structure of the Universe and of superfluid ^3He go further. Not only is the symmetry of the coherent quantum state of the superfluid analogous to the quantum vacuum of the Universe, but the excitations and collective modes of the superfluid have analogies with the fundamental particles bosons and fermions (ref. 4). Finally, various types of linear defects (vortices), point defects (monopoles) and textures may be generated in superfluid ^3He , in analogy with the various types of defects proposed for the structure of Universe.

The Kibble mechanism for vortex creation during phase transitions¹ is illustrated schematically for our experiment in Fig. 1. After a sudden local heating, the ^3He recools through the transition to the superfluid state. During this process, separate superfluid regions are independently nucleated with random orientations of the order parameter in each domain. As the domains grow and make contact with their neighbours, the resulting superfluid cannot be uniform. The subsequent order-parameter 'glass' forces a distribution of topological defects leading to a tangle of quantized vortex lines.

The size of the initial domains depends strongly on the rapidity with which the transition is traversed. According to Zurek³ the distance between the ensuing vortices β is given approximately by $\beta = \xi_0(\tau_Q/\tau_0)^{1/4}$, where ξ_0 is the coherence length at zero temperature, τ_0 , the coherence length divided by the Fermi velocity, (ξ_0/v_F) , is the characteristic time of the superfluid, and τ_Q is the characteristic time for cooling through the phase transition.

Superfluid ^3He has the further advantage as the working substance in that it allows a very precise localized crossing of the phase transition to be generated by nuclear reaction⁵, which avoids the more violent global processes used for earlier experiments on ^4He and liquid crystals^{2,6}. The ^3He nucleus has a very high cross-section for neutrons via the process: $n + ^3\text{He} \rightarrow ^3\text{H} + p$. This process also liberates a precise energy of 764 keV which is initially shared by the product proton and tritium nucleus. The mean free path of the products in the liquid is limited to around 30 μm as the kinetic energy is rapidly thermalized with the creation of a cloud of quasiparticle excitations and excited ^3He atoms providing enough heat to warm a small volume of the liquid to above the superfluid transition. The volume of normal liquid ^3He then rapidly recools on a timescale of 1 μs or less.

In a companion experiment⁷ Ruutu *et al.* have shown directly that vortices are indeed created in superfluid ^3He by this process. They observe unambiguous vortex nucleation triggered by neutron irradiation in a rotating experiment at higher temperatures. In the present experiment we measure how much of the energy released is retained in the liquid in the form of a vortex tangle. As we know the energy of a vortex per unit length, we can go further and make a quantitative calculation of the vortex density remaining which we can compare with Zurek's prediction.

The present measurements are made close to the lower limit to

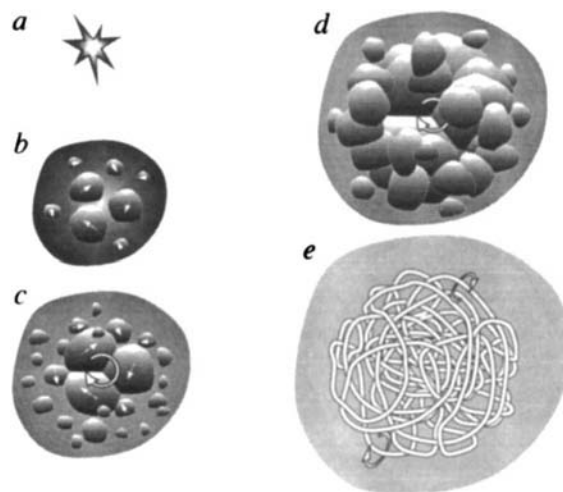


FIG. 1 Schematic view of the creation of a tangle of linear singularities by the nucleation of independent ordered regions as a system passes through a phase transition. In the superfluid ^3He context; at a, a neutron has struck a ^3He nucleus initiating the process $n + ^3\text{He} \rightarrow ^3\text{H} + p$, liberating 764 keV of energy and creating a small region of very high temperature in the liquid. At a later time b the hot region has expanded and cooled to near the transition temperature. Small regions of superfluid are independently nucleated each having a different value of the order parameter, as shown by the small arrows. At c the three central ordered regions now touch. Although the order parameter may bend to accommodate the boundaries, a full 2π change around the triple contact point remains. This is a vortex. At d, many more regions have been nucleated and overlap, and along the grain boundaries a whole tangle of vortex lines is created. Finally, at e the hot bubble has cooled entirely through the transition and only the tangle of vortices remains.

which superfluid ^3He can be cooled (around $0.1 T_c$, where T_c is the superfluid transition temperature). At such temperatures the few remaining thermal quasiparticles form an extremely dilute non-interacting gas. It is the mutual friction provided by the quasiparticles which mediates the evolution and gradual disappearance of vortices. Therefore, under our conditions, after a neutron event, the local high density of quasiparticles falls very rapidly to the vanishingly small background density. Thus the vortices are only exposed to appreciable mutual friction for a very short time and we estimate that almost all the initial vortex density should survive.

The experiment is made at temperatures down to $160\text{ }\mu\text{K}$ in the Grenoble nuclear refrigerator. The detector is a Lancaster-style quasiparticle detector⁸, in the form of a 0.15 cm^3 copper-walled enclosure containing superfluid ^3He , connected via a $60\text{-}\mu\text{m}$ hole to a larger volume of superfluid at a temperature below $100\text{ }\mu\text{K}$. Energy deposited within the enclosure warms the superfluid by the creation of quasiparticle excitations. The resulting temperature rise is detected by a sensitive vibrating-wire resonator. The temperature of the enclosure then recovers as quasiparticles escape through the small hole over a time determined by the geometry. A particle scattering event inside the detector is observed on the temperature record as a sharp temperature rise followed by a slow recovery over about 60 s.

The energy sensitivity of the detector is calibrated with a second vibrating-wire resonator which serves as a heater (which, when intensely oscillated, creates quasiparticle-quasihole pairs in the superfluid). We can simulate a scattering event by exciting the 'heater' resonator to introduce a measured quantity of energy during a time which is short compared to the detector recovery time, allowing a rather accurate energy calibration of the detector.

An external AmBe (americium-beryllium) neutron source irradiates the cell with a very low flux of neutrons. The temperature trace is recorded, and the events analysed and presented as a spectrum of counts versus energy. The whole procedure is repeated at three different pressures for the superfluid (because the superfluid properties, such as gap and coherence length, are pressure-dependent). We observe a clear peak in the energy spectra associated with the $n + ^3\text{He} \rightarrow p + ^3\text{H} + 764\text{ keV}$ reaction, as shown in Fig. 2. Significantly, however, the peak occurs at an energy substantially below the canonical 764 keV. Clearly not all the energy deposited in the detector by the reaction appears as thermal excitations.

Apart from appearing as thermal excitations, the deposited energy can take two other paths. First, the energy can appear as ultraviolet photons. It is well known that collision events in liquid ^4He result in scintillation. The ^4He ions produced after the collision form dimers which decay by a cascade of processes ultimately leading to the emission of ultraviolet photons to which helium is transparent. These photons are absorbed in the surrounding walls and their energy is prevented from returning to the liquid helium by the large thermal boundary resistance. The ultraviolet photon process⁹ accounts for 6–8% of the total energy deposition in ^4He and is not likely to be at all sensitive to which isotope is involved. The second path involves the creation of defects in the superfluid, of which only vortices can acquire a significant energy.

As we can be certain that the ultraviolet losses are very insensitive to pressure, we can subtract the ultraviolet component from the measured energy missing from the initial 764 keV. We find that the remaining energy still unaccounted for is 60, 70 and 125 keV at the three pressures measured, 0, 6 and 19.4 bar, respectively. We believe that this remaining energy deficit has been retained in the liquid by the creation of vortices.

To calculate the initial separation of the vortices, β , we need to know, first, the volume of liquid ^3He heated above the transition, second, the energy per unit length of the vortex, and third, for comparison with Zurek's predictions, we also need an estimate of the time for cooling through the transition, τ_Q . The energy per unit length of a vortex is approximately $(\rho/4\pi)(\hbar/2m_3)^2 \ln(\beta/\xi_0)$ which

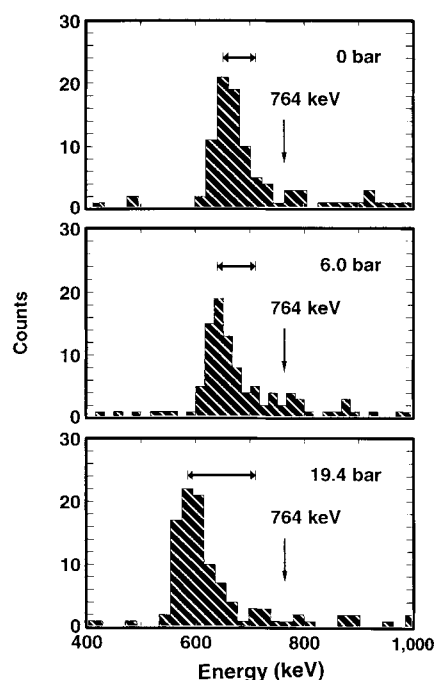


FIG. 2 The energy spectra in counts versus energy deposited into the quasiparticle system of superfluid ^3He under neutron irradiation at three pressures of the liquid. The peak from the neutron-capture process $n + ^3\text{He} \rightarrow ^3\text{H} + p$ is clearly visible. This reaction is exothermic, liberating 764 keV of energy. Approximately 7% of this energy is released as ultraviolet photons which are lost from the helium. However, as shown in the figure by the double-ended arrow, a significant fraction of energy is 'missing', a quantity which increases with pressure. It is this energy, which does not produce thermal excitations, that we assume remains in the liquid in the form of a vortex tangle.

is about 0.5 keV mm^{-1} at 0 bar. The quantities V and τ_Q are sensitive to the details of the process whereby the kinetic energy of the released proton and tritium nucleus is converted to excitations of the fluid. This process is very poorly understood. However, we can make estimates with the following very simple model: the energy E is assumed to be deposited at a point in the fluid, which subsequently cools by thermal diffusion with the diffusion constant D taken to be equal to that of the ^3He liquid at the superfluid transition temperature T_c . These assumptions yield a spherical volume of normal fluid with radius $R \approx 0.4(E/CT_c)^{1/3}$, where C is the liquid heat capacity just above T_c and $E = 710\text{ keV}$ is the deposited energy after subtraction of the ultraviolet losses. The characteristic cooling time is then given by $\tau_Q \approx R^2/4D$. This yields normal volumes with radii of 25, 16 and $11\text{ }\mu\text{m}$ and cooling times of 0.23, 0.46 and $1.1\text{ }\mu\text{s}$ for the measured 0, 6 and 19.4 bar pressures, respectively.

The coherence length in superfluid ^3He for the three pressures considered here is 770, 390 and $210\text{ }\text{\AA}$. Our measured value of the ratio β/ξ_0 , (the separation of the vortices in coherence-length units) is thus found to have a pressure independent value of about 8. We can compare this ratio with the appropriate value of the equivalent parameter $(\tau_Q/\tau_0)^{1/4}$ in the Zurek model. Using the simple model of diffusion cooling for the hot volume considered above, we find a value of this parameter varying from 4 to 7 over the pressure range investigated.

Given the simplicity of the assumptions and approximations used, the agreement between our measured values of β/ξ_0 and the corresponding Zurek parameter, $(\tau_Q/\tau_0)^{1/4}$, is surprisingly close. Because the whole process of the conversion of the initial reaction energy into thermal energy is at present poorly understood, it is very difficult to make an accurate estimate of the errors involved

in the simple model we have used for the cooling process. Furthermore, the prediction $\beta/\xi_0 = (\tau_Q/\tau_0)^{1/4}$ from the Zurek model is only qualitative.

This close agreement, however, does confirm that a fast cooling through the superfluid transition indeed creates a residual density of vortices as foreseen in Zurek's model of the Kibble mechanism. That the observed vortices are further spaced than the theory suggests may simply be a reflection of the approximations made both in deducing the value from the data and estimating the equivalent number from Zurek's expression. Why the observed separation, β/ξ_0 , is apparently independent of pressure is not clear at present. Finally, if indeed analogous topological objects have once existed or even survive in the Universe, then we have taken the first tentative steps with superfluid ^3He towards the quantitative modelling of their formation. $\square$

Received 15 January; accepted by 4 June 1996.

1. Kibble, T. W. B. *J. Phys.* **A9**, 1387–1398 (1976).
2. Hendry, P. C., Lawson, N. S., Lee, R. A. M., McClintock, P. V. E. & Williams, C. H. D. *Nature* **368**, 315–317 (1994).
3. Zurek, W. H. *Nature* **317**, 505–508 (1985).
4. Volovik, G. E. *Exotic Properties of Superfluid ^3He* (World Scientific, Singapore, 1992).
5. Bradley, D. I. et al. *Phys. Rev. Lett.* **75**, 1887–1891 (1995).
6. Chuang, I., Durrer, R., Turok, N. & Yurke, B. *Science* **251**, 1336–1342 (1991).
7. Ruutu, V. M. H. et al. *Nature* **382**, 334–336 (1986).
8. Fisher, S. N., Guénault, A. M., Kennedy, C. J. & Pickett, G. R. *Phys. Rev. Lett.* **69**, 1073–1076 (1992).
9. Adams, J. S. et al. *Phys. Lett.* **B341**, 431–434 (1995).

ACKNOWLEDGEMENTS. We thank G. Volovik for discussions, and J.-L. Bret, J. Chaussy and L. Puech for their assistance with the electronics and computer programming. This work was supported in part by Etat/DRET.

CORRESPONDENCE should be addressed to Y.M.B. (e-mail: bunkov@labs.polycnrs-gre.fr).

Vortex formation in neutron-irradiated superfluid ^3He as an analogue of cosmological defect formation

V. M. H. Ruutu*, V. B. Eltsov*†, A. J. Gill‡§, T. W. B. Kibble‡, M. Krusius*, Yu. G. Makhlin*||, B. Plaçais¶, G. E. Volovik*|| & Wen Xu*

* Low Temperature Laboratory, Helsinki University of Technology, 02150 Espoo, Finland

† Kapitza Institute for Physical Problems, 117334 Moscow, Russia

‡ Blackett Laboratory, Imperial College, London SW7 2BZ, UK

§ T-6 Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

|| Landau Institute for Theoretical Physics, 117334 Moscow, Russia

¶ Laboratoire de Physique de la Matière Condensée de l'Ecole Normale Supérieure, CNRS URA 1437, F-75231 Paris Cedex 05, France

TOPOLOGICAL defects formed during a rapid symmetry-breaking phase transition in the early Universe^{1,2} could be responsible for seeding large-scale structure, for the anisotropy of the microwave background radiation, and for the predominance of matter over antimatter^{3,4}. The theory describing this cosmological phase transition is formally analogous to that describing the transition to the superfluid state in liquid ^3He , so that in principle the process of cosmological defect formation can be modelled in the laboratory. Here we report the results of an experiment in which the 'primordial fireball' is mimicked using a neutron-induced nuclear reaction ($n + ^3\text{He} \rightarrow p + ^3\text{He} + 0.76 \text{ MeV}$) to heat small regions of superfluid ^3He above the superfluid transition temperature. These bubbles of normal liquid cool extremely rapidly, and we find that their transition back to the superfluid state is accompanied by the formation of a random network of vortices

(the superfluid analogue of cosmic strings). We monitor the evolution of this defect state by rotating the superfluid sample, allowing vortices to escape from the network and thus be probed individually. Our results provide clear confirmation of the idea that topological defects form at a rapid second-order phase transition, and give quantitative support to the Kibble–Zurek mechanism^{5,6} of cosmological defect formation.

Laboratory experiments intended to test different theories of cosmological defect formation have recently been conducted in liquid crystals^{7,8} and superfluid ^4He (ref. 9). Use of the superfluid phases of liquid ^3He allows further progress in these Big Bang simulations. ^3He has several advantages over other systems. Many direct parallels and formal analogies connect superfluid ^3He theory with the field theories that are used to describe the physical vacuum, gauge fields and fermionic elementary particles¹⁰. Superfluid ^3He exhibits a variety of phase transitions and topological defects which can be detected with NMR methods with single-defect sensitivity¹¹. Different modern models of defect-mediated baryogenesis^{3,4,12} can be simulated by the interaction of the ^3He defects, vortices and domain walls, with fermionic quasiparticles. Of particular relevance to our rapid-quench experiment is the fact that liquid ^3He can be efficiently locally heated with thermal neutrons (Fig. 1).

The neutrons are produced with a paraffin-moderated Am–Be source of 10 mCi activity. They are then incident upon the ^3He sample container. At the minimum distance of 22 cm between

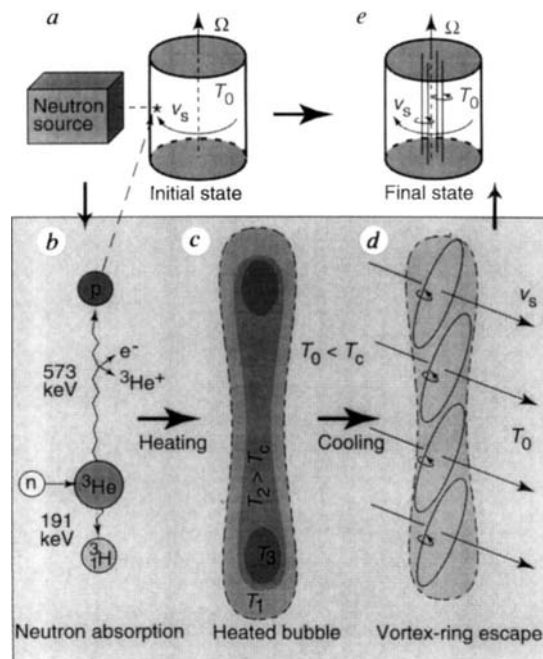


FIG. 1 Principle of the experiment and of vortex formation during neutron irradiation: *a*, Superfluid $^3\text{He-B}$ contained in a cylinder, which rotates at constant angular velocity Ω , is exposed to neutron radiation. *b*, In a neutron-absorption event, a proton and a tritium nucleus are formed which fly apart in opposite directions. Their kinetic energies are used up in ionization reactions of ^3He atoms. *c*, The ionized particles recombine, generating heat in the form of quasiparticle excitations which corresponds to a substantial part of the total reaction energy. A cigar-shaped hot region of liquid is formed where the temperature rises above T_c . *d*, The hot bubble cools rapidly towards the bulk liquid temperature T_0 . While passing through the superfluid transition at T_c a tangled network of vortex lines is formed. Vortex loops which exceed a critical size start to expand in the superfluid flow created by the rotation of the container whereas smaller loops contract and annihilate. *e*, An expanding vortex ring ultimately intersects the walls of the container and gives rise to one rectilinear vortex line. The vortex lines accumulate in the centre of the rotating container where they are counted with NMR.

source and sample, $v \approx 20$ neutrons min^{-1} are absorbed by the ^3He liquid. For thermal neutrons the mean free path is 0.1 mm in liquid ^3He , and thus all absorption reactions occur close to the walls of the cylindrical container which has radius $R = 2.5$ mm. Because the incident thermal neutron has low momentum, the 573-keV proton and 191-keV tritium nucleus fly apart in opposite directions, producing ionization tracks that are 70 and 10 μm long, respectively. The subsequent charge recombination yields a heated region of the normal liquid phase which, for simplicity, we here assume to be of a spherically symmetric shape¹³.

The bubble of normal fluid cools by the diffusion of quasi-particle excitations out into the surrounding superfluid with a diffusion constant $D \approx v_F l$ where v_F is their Fermi velocity and l the mean free path. The difference from the surrounding bulk temperature T_0 as a function of the radial distance r from the centre of the bubble is given by $T(r, t) - T_0 \approx E_0 \exp(-r^2/4Dt)/(C_v(4\pi Dt)^{3/2})$ where E_0 is the energy deposited by the neutron as heat and C_v is the specific heat. The maximum value of the bubble radius R_b , with fluid in the normal phase above the superfluid transition temperature $T_c \approx 2$ mK, is given for $T(r) > T_c$ by

$$R_b \approx (E_0/C_v T_c)^{1/3} (1 - T_0/T_c)^{-1/3} \quad (1)$$

which is typically of the order of 10 μm . The bubble cools and shrinks away rapidly with the characteristic time $\tau_Q \approx R_b^2/D \approx 10^{-6}$ s. The rapid superfluid phase transition leads to the formation of a random network of vortices. The fate of vorticity depends on the bias field, the external superfluid velocity $v_s = \Omega R$, created by rotating the container at an angular velocity Ω . If the radius of the vortex loop exceeds the value $r_0(v_s) = (\kappa/4\pi v_s) \ln r_0/\xi$, where $\kappa = \pi\hbar/m_3$ is the circulation quantum and $\xi = \xi_0(1 - T/T_c)^{-1/2}$ is the superfluid coherence length, the loop expands. An expanding vortex ring eventually results in a rectilinear vortex line which is pulled to the centre of the container. If several vortex lines are formed they accumulate in a central vortex bundle¹¹.

The number of the vortex lines is monitored with NMR. In Fig. 2 inset two NMR absorption records are shown, starting from the moment when the neutron source is placed in position. The absorption events, which lead to vortex formation, are visible as distinct steps. The step height gives the number of vortex lines created per event, if the neutron source is sufficiently far from the sample and the absorption events are well separated in time. The number of vortex lines created per unit time increases rapidly with increasing v_s (Fig. 2). The extrapolation to zero gives the threshold value v_{cn} plotted in Fig. 3 as a function of T_0 for different pressures. This v_{cn} is much smaller and has a different dependence on T_0 than the critical velocity v_c at which a vortex is formed in the absence of neutrons in the B phase of superfluid ^3He ($^3\text{He-B}$)¹¹.

According to ref. 2, the topological defects are created during the symmetry-breaking transition when the order parameter begins to fall from the false ground state into the true degenerate ground states which form independently in regions which are not causally connected. The different regions grow in size, and ultimately the order parameter fills all space but with a multitude of defects. The important notion in this theory is the initial density of defects, that is, the defect density at the moment when the defects can be distinguished from the background fluctuations of the order parameter. In the original Kibble model² this was identified with the moment when the system cools below the Ginzburg temperature T_G above which the thermal fluctuations prevail. This implies that the initial distance between the vortices is the coherence length at T_G , that is, $\xi(T_G) = \xi_0/\sqrt{1 - (T_G/T_c)}$. In superfluid ^3He the critical fluctuation region is extremely narrow, $1 - (T_G/T_c) \approx (a/\xi_0)^4$, because the zero temperature coherence length $\xi_0 \approx 0.1 \mu\text{m}$ is two orders of magnitude larger than the interatomic spacing a (ref. 14). As a result $\xi(T_G) \approx \xi_0(\xi_0/a)^2$ exceeds the bubble size R_b , which means that no vortices can be created in this model.

In an alternative theory put forward by Zurek^{5,6}, the vortices are formed at the moment when the coherence length becomes

comparable with the causal horizon. For ^3He this occurs at the time $\sqrt{\tau_Q \tau_0}$ after the transition, where $\tau_0 = \xi_0/v_F$. The value of the coherence length ξ at that instant, $\xi_0(\tau_Q/\tau_0)^{1/4}$, gives the initial distance between the defects, of the order of 1 μm in our case. This is well within R_b , in agreement with our experiment.

The later evolution of the vortex network leads to a gradual increase of the intervortex distance $\xi(t)$, while the network structure remains scale-invariant^{3,4}. For a homogeneous system the number of loops $n(l)$ per unit length and unit volume with line lengths $l > \xi$ is given¹⁵ by $n(l) = C \xi^{-3/2} l^{-5/2}$ where $C \approx 0.4$. For our case of finite bubble volume, where all vortices form closed loops, our numerical simulation gives the same distribution law but with slightly different $C \approx 0.3$. For the average straight-line dimension $\mathcal{D}$ of a loop we find $\mathcal{D} = \beta(l\xi)^{1/2}$ with $\beta \approx 1$. The loop size distribution in terms of $\mathcal{D}$ is $n(\mathcal{D}) d\mathcal{D} \approx 2C d\mathcal{D}/\mathcal{D}^4$, and does not depend on the distance between vortices. The evolution of the network leads to an increasing lower cut-off of the distribution, $\mathcal{D}_{\min} = \alpha \xi(t)$, where α is of the order of unity. The upper cut-off is the diameter of the bubble $\mathcal{D}_{\max} = 2R_b$.

When the average radius of curvature, determined by ξ , exceeds $r_0(v_s)$ the vortices start to escape from the bubble. The number of vortices, $\mathcal{N}(v_s)$, created per neutron is thus the number of loops with $\alpha r_0(v_s) < \mathcal{D} < 2R_b$ in the bubble volume V_b :

$$\mathcal{N}(v_s) = V_b \int_{\alpha r_0(v_s)}^{2R_b} d\mathcal{D} n(\mathcal{D}) = \frac{\pi C}{9} \left[\left(\frac{2R_b}{\alpha r_0(v_s)} \right)^3 - 1 \right] \quad (2)$$

The critical velocity v_{cn} is determined by the requirement $\mathcal{N}(v_{cn}) = 0$, which gives $v_{cn} = (\kappa\alpha/8\pi R_b) \ln(R_b/\xi)$. Our numerical simulation suggests $\alpha \approx 1$. Thus v_{cn} is inversely proportional to R_b ,

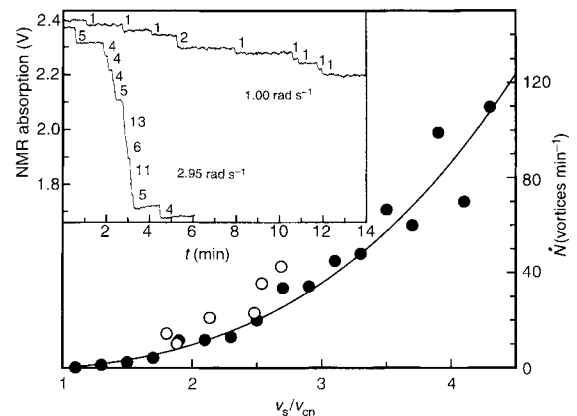


FIG. 2 Formation of vortices during neutron irradiation. Inset, change in NMR absorption as a function of running time t at high and low rotation velocity. The neutron source is turned on at $t = 0$. Each step corresponds to one neutron absorption event. The height of the step gives the number of vortex lines created and is denoted by the adjacent number. The high-velocity trace is recorded with the neutron source at a 2.8 times larger distance from the $^3\text{He-B}$ sample. The bulk liquid temperature is $T_0 = 0.96 T_c$. Main figure, the rate N at which vortex lines are created during neutron irradiation, plotted as a function of the normalized superfluid velocity v_s/v_{cn} . Below the critical threshold at v_{cn} , the rate vanishes, while above the threshold the rate follows the fitted cubic dependence $N = \gamma[(v_s/v_{cn})^3 - 1]$, with $\gamma = (1.37 \pm 0.06)$ vortex lines per min, shown by the solid line. The rate of vortex formation is the number of vortices per neutron $\mathcal{N}(v_s)$ multiplied by the neutron flux v : $\dot{N} = v \mathcal{N}(v_s)$. The theoretical estimate for $\mathcal{N}(v_s)$ from equation (3) with $v \approx 20$ neutrons min^{-1} gives $\gamma = v\pi C/9 \approx 2$, which is in order-of-magnitude agreement with the experiment. The filled circles show data measured at a bulk liquid temperature $T_0 = 0.96 T_c$; the open circles, data measured at $0.90 T_c$. Both sets of data fall on the same cubic dependence, although v_{cn} differs by 50% (Fig. 3). The cubic dependence on the superflow velocity is ultimately interrupted by the spontaneous critical velocity $v_c(T_0, P)$, which is the maximum possible superflow velocity and here at $0.96 T_c$ about $6.5 v_{cn}$. The liquid pressure $P = 18.0$ bar and the magnetic field $B = 11.8$ mT are the same as in the inset.

and according to equation (1) has the temperature dependence $v_{cn} \propto (1 - T_0/T_c)^{1/3}$, in agreement with the measurements in Fig. 3. In terms of v_{cn} one obtains the universal curve

$$\mathcal{N}(v_s/v_{cn}) = \frac{\pi C}{9} \left[\left(\frac{v_s}{v_{cn}} \right)^3 - 1 \right] \quad (3)$$

Equation (3) reproduces the observed cubic velocity dependence and gives the correct order of magnitude estimate of $\mathcal{N}(v_s)$, as seen in Fig. 2. It also carries the same universality feature as the measured results in Fig. 3: the ambient measuring conditions depend on what values are chosen of the temperature T_0 , pressure P , and magnetic field B , but in the results for $\mathcal{N}(v_s)$ all this dependence is contained in the single parameter v_{cn} .

The experiment demonstrates that vortices are created within bulk liquid ^3He during a rapid quench to the superfluid state. The measured number of vortices created as a function of velocity is consistent with the Vachaspati–Vilenkin scaling law¹⁵ discussed for cosmic strings. This is a strong argument in favour of the Kibble–Zurek mechanism of defect formation. The external flow mainly plays the role of the bias field which allows vortices to escape, to expand and finally to reach a stable state in the rotating container so that they can be detected one-by-one. With increasing superfluid velocity v_s , smaller loops, which represent an earlier stage in the evolution of the network, are extracted. We expect that when we reach the smallest possible size of loops, which is limited by the initial interdefect distance, the cubic scaling law will be violated. At the moment, the smallest size of the loops extracted at the highest rotation velocity is above, but close to, the limit which follows from the Zurek model, but is several orders of magnitude smaller than suggested by the original Kibble theory. This supports the Zurek modification of the Kibble mechanism. We are planning to study the Kibble mechanism for formation of other types of defects in superfluid ^3He , such as half-quantum vortices (Alice strings) and domain walls. Superfluid ^3He

is also the only condensed-matter system where one can try to simulate the processes of interaction of defects with fermions that lead to the baryon asymmetry.

The Kibble–Zurek mechanism, based on a simple physical picture, gives a good quantitative fit to our measurements. Vortex formation in neutron-irradiated superfluid ^3He now provides the most stringent confirmation of the formation of topological defects at a rapid second-order phase transition, and allows quantitative comparison with the estimated loop size distribution. This gives confidence in applications of similar ideas to the early Universe, where topological defects remain a viable explanation for large-scale structure formation and the microwave background anisotropy. Detailed predictions, for example of the intensity of gravitational radiation emitted by oscillating cosmic strings, are strongly dependent on the initial density of strings at formation. $\square$

Received 21 December 1995; accepted 4 June 1996.

1. Zeldovich, Ya. B., Kobzarev, I. Yu. & Okun, L. B. *Zh. éksp. teor. Fiz.* **67**, 3–11 (1974); *JETP* **40**, 1–5 (1974).
2. Kibble, T. W. B. *J. Phys.* **A9**, 1387–1398 (1976).
3. Vilenkin, A. & Shellard, E. P. S. *Cosmic Strings and Other Topological Defects* (Cambridge Univ. Press, 1994).
4. Hindmarsh, M. B. & Kibble, T. W. B. *Rep. Prog. Phys.* **58**, 477–562 (1995).
5. Zurek, W. H. *Nature* **317**, 505–508 (1985).
6. Zurek, W. H. *Los Alamos preprint LA-UR-95-2269* (1995); *Phys. Rep.* (in the press); *Acta physica polonica* **B24**, 1301–1311 (1993).
7. Chuang, I., Durrer, R., Turok, N. & Yurke, B. *Science* **251**, 1336–1342 (1991).
8. Bowick, M. J., Chander, L., Schiff, E. A. & Srivastava, A. M. *Science* **263**, 943–945 (1994).
9. Hendry, P. C., Lawson, N. S., Lee, R. A. M., McClintock, P. V. E. & Williams, C. D. H. *Nature* **368**, 315–317 (1994).
10. Volovik, G. E. & Vachaspati, T. *Int. J. mod. Phys.* **B10**, 471–521 (1996).
11. Parts, Ü et al. *Europhys. Lett.* **31**, 449–454 (1995).
12. Turok, N. in *Formation and Interaction of Topological Defects* (eds Davis, A. C. & Brandenberger, R.) 283–301 (Plenum, New York, 1995).
13. Schiffer, P. & Osheroff, D. D. *Rev. mod. Phys.* **67**, 491–501 (1995).
14. Vollhardt, D. & Wölfle, P. *The Superfluid Phases of ^3He* (Taylor & Francis, London, 1990).
15. Vachaspati, T. & Vilenkin, A. *Phys. Rev.* **D30**, 2036–2045 (1984).

ACKNOWLEDGEMENTS. This collaboration was carried out under the EU Human Capital and Mobility programme. A.J.G., T.W.B.K. and B.P. thank the Low Temperature Laboratory for support and hospitality.

CORRESPONDENCE should be addressed to T.W.B.K. (e-mail: kibble@ic.ac.uk).

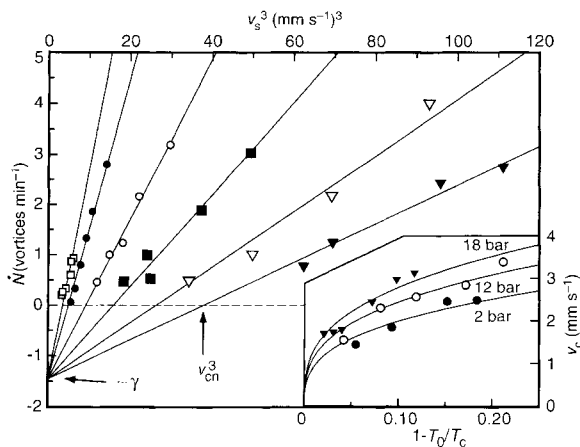


FIG. 3 Inset, critical superflow velocities as a function of the normalized bulk liquid temperature $1 - T_0/T_c$ and pressure P of the liquid ^3He -B sample in a magnetic field of $B = 11.8\text{ mT}$. Solid lines, critical value of superflow velocity $v_{cn} \propto (1 - T_0/T_c)^{1/3}$ at which vortex creation starts in the presence of the neutron source. Main figure, the rate $\dot{N}$ at which vortex lines are created during neutron irradiation, plotted as a function of v_s^3 , demonstrating the scale-invariance of the vortex-formation law in different ambient conditions: $P = 2\text{ bar}$, $T_0 = 0.95T_c$, $B = 11.8\text{ mT}$ (empty squares); $P = 12\text{ bar}$, $T_0 = 0.96T_c$, $B = 11.8\text{ mT}$ (filled circles); $P = 18\text{ bar}$, $T_0 = 0.97T_c$, $B = 16\text{ mT}$ (empty circles); $P = 18\text{ bar}$, $T_0 = 0.97T_c$, $B = 28\text{ mT}$ (filled squares); $P = 18\text{ bar}$, $T_0 = 0.90T_c$, $B = 28\text{ mT}$ (empty triangles); $P = 18\text{ bar}$, $T_0 = 0.91T_c$, $B = 60\text{ mT}$ (filled triangles). In agreement with equation (3), the dependence on temperature T_0 , pressure P , and magnetic field B is contained in the critical velocity v_{cn} . The intercept of the fitted lines with the horizontal axis yields v_{cn}^3 . The common intercept with the vertical axis, $\gamma = (1.46 \pm 0.12)$ vortex lines per min, defines the numerical prefactor in equation (3).

An explanation for the central stress minimum in sand piles

J. P. Wittmer*, P. Claudin†, M. E. Cates* & J.-P. Bouchaud‡

* Department of Physics and Astronomy, University of Edinburgh, JCMB King's Buildings, Mayfield Road, Edinburgh EH9 3JZ, UK

† Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, UK

‡ Service de Physique de l'Etat Condensé, CEA Ormes des Merisiers, 91191 Gif-sur-Yvette, Cedex France

A KEY component of a continuum-mechanical description of a sand pile (viewed as an assembly of hard particles in frictional contact) is the requirement of stress continuity—the forces acting on a small element of material must balance. But an additional physical postulate is required to close the equations. Standard continuum approaches postulate that the material is everywhere just at the point of slip failure^{1,2}. But these approaches have been unable to explain a startling experimental observation³—that the weight exerted by a conical sand pile on a surface has a minimum, not a maximum, below the apex. Here we propose a new closure, which embodies an intuitive model of arching^{4,5} within a fully consistent continuum theory. Our assumption is that the principal stress axes have a fixed angle of inclination to the vertical. In two dimensions, this is sufficient to close the equations. In three dimensions, a second closure

relation is required, but our results are relatively insensitive to the choice made. Our model, which contains no adjustable parameters, can account for the vertical stress distribution in real sand piles. This supports the idea that stresses propagate within a granular medium according to local rules that depend on its construction history.

The equations of stress continuity express the fact that, in static equilibrium, the forces acting on a small element of material must balance. For a pile of sand we have, in $d = 3$ dimensions,

$$\partial_r \sigma_{rr} + \partial_z \sigma_{rz} = \alpha(\sigma_{zz} - \sigma_{rr})/r \quad (1a)$$

$$\partial_r \sigma_{rz} + \partial_z \sigma_{zz} = g - \alpha \sigma_{zz}/r \quad (1b)$$

$$\partial_\chi \sigma_{ij} = 0 \quad (1c)$$

where $\alpha = 1$, and z, r and χ are cylindrical polar coordinates, with z the downward vertical. We take $r = 0$ as a symmetry axis, so that $\sigma_{r\chi} = \sigma_{z\chi} = 0$; g is the force of gravity per unit volume; σ_{ij} is the usual stress tensor which is symmetric. The equations for $d = 2$ are obtained by setting $\alpha = 0$ in equation (1a, b) and suppressing equation (1c).

In two dimensions (we return to $d = 3$ later) one has two equations in three unknowns, and one further equation is required. This is the closure problem of granular mechanics which, despite a century of study^{1,2}, remains open. In an elastic material, the missing equation is the constitutive relation (c.r.) between stress and strain. For a cohesionless granular medium, made of hard particles in frictional contact, no 'strain' variable can be defined; we should instead seek a c.r. (which we assume to be local) between one stress component and another. This assumes that the slight deformability of real grains can be ignored (see below).

Much of the literature^{1,2} assumes as an (implicit) c.r. 'incipient failure everywhere' (IFE)⁶: through every point in the material there passes a plane across which the shear stress is the maximum allowed by static friction. Thus, each material element is just at the point of slip failure. However, for a sand pile at its angle of repose, ϕ , this incipient failure condition can safely be assumed only at the surface of the pile (if we ignore some small hysteresis effects^{7,8}), and not in the bulk. Moreover, the IFE closure does not explain a

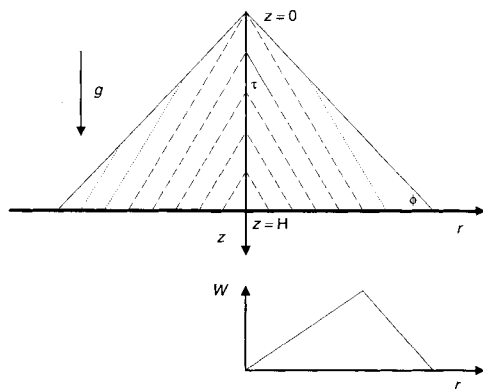


FIG. 1 The Edwards arch picture, and the resulting distribution of the vertical stress. The weight, W , of the sand is viewed as transmitted in straight lines down one of a series of nested arches; points at the base support a weight proportional to the length of the arch above. The angle τ is not fixed within the model; $W = 0$ on the centre line (overpredicting the dip). Note that the outer 'arches' are subject to unbalanced forces in this scalar model. In the FPA model, the arch direction becomes the major principal stress axis; the angle τ can then be shown to obey $\tau = (\pi - 2\phi)/4$ (where ϕ is the repose angle), using the stress continuity equations and the condition of incipient failure at the free surface. (The latter is the Coulomb yield condition, $|\sigma_{ns}| = \sigma_{nn} \tan(\phi)$, where n and s are directions normal and parallel to the surface.) The angle of the arch changes discontinuously at $r = 0$, giving a non-analytic behaviour there.

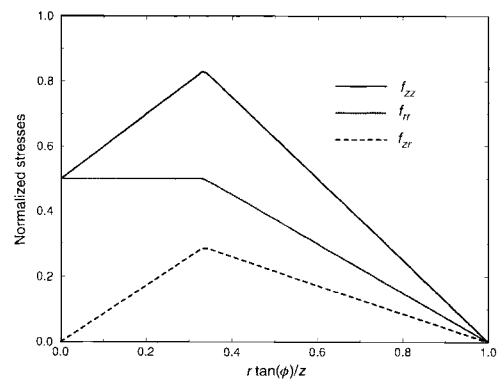


FIG. 2 The distribution of shear and vertical stresses in a two-dimensional sand-pile, found using the FPA closure (for $\phi = 30^\circ$). The behaviour of σ_{zz} can be established by substituting equation (1a) and equation (2) into equation (1b), which gives a wave equation $(\partial_z - c_1 \partial_r)(\partial_z - c_2 \partial_r)\sigma_{zz} = 0$, with $c_1 + c_2 = -2 \tan \phi$ and $c_1 c_2 = -1$. (A similar equation applied to σ_{rr} .) The characteristics of this equation are two rays, one at angle τ to the vertical (down the 'arch' direction) and the other at right angles. (The BCC model⁶ instead has characteristics symmetric about the vertical. The solution method is similar, except that the rays now suffer both a reflection and a refraction at the centre line.) The maximum vertical stress occurs at $r/z = \tan \tau$, and is a factor $(1 + 2 \tan^2 \phi)$ times larger than the value at $r = 0$. In two dimensions, incipient failure is present (though not assumed) for all $r > z \tan \tau$; in three dimensions, however, it occurs only at the free surface. (Note that Trollope's model^{14,15} resembles BCC but introduces *ad hoc* a third ray that apparently cannot be present if the c.r. is local. Note also that the FPA closure can be viewed as a limiting case of a class of nonlinear models studied in section 3.3 of ref. 6.) f_{zz} , f_{rr} and f_{rz} are respectively the vertical, radial and shear components of the reduced stress tensor $f_{ij} = \sigma_{ij}/gz$.

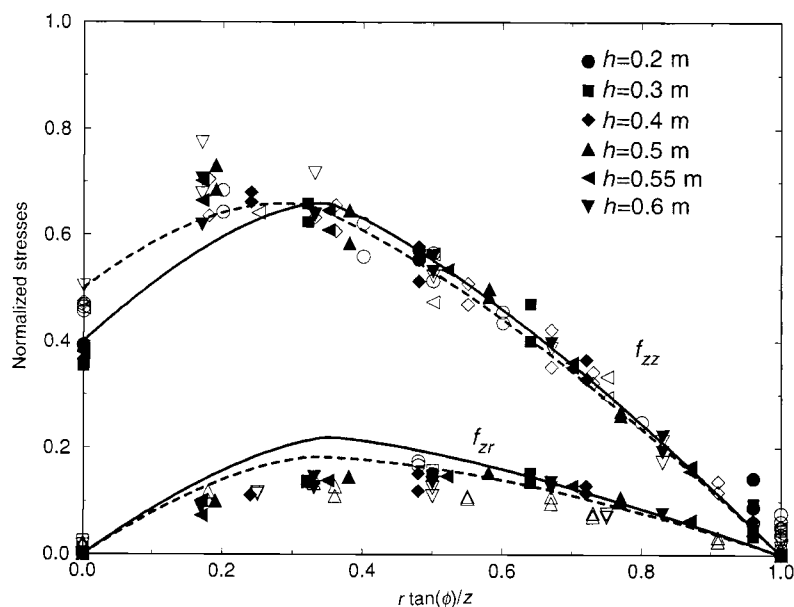
dramatic experimental fact³: the vertical stress $\sigma_{zz}(r)$ beneath a sand pile has a minimum ('the dip') at the centre^{3,9} (seen also in some simulations¹⁰). The IFE closure, and other proposals⁶, give either a smooth maximum or flat plateau⁶, as do related discrete models^{11,12}. For all these models (but not for some others¹³), the stress can be written $\sigma_{ij}(r, z) = gzf_{ij}(r \tan \phi/z)$ (where f_{ij} , the reduced stress, vanishes on the free surface). This 'radial stress-field' (RSF) scaling^{1,2} is obeyed by experiment (see below); it means there is no intrinsic length scale in the problem of a sand-pile under gravity. A finite elastic modulus Y for the grains introduces a length Y/g ; the observation of RSF scaling suggests that deformability effects are negligible.

A strongly contrasting approach is that of Edwards^{4,5}, who gave a simple scalar argument for the dip (Fig. 1) based on the idea of arches. (Similar ideas can be found in the earlier work of Trollope^{14,15}.) Despite its appeal, Edwards' approach is inconsistent in attempting to find the normal stress while ignoring shear stresses: the model is mechanically unstable (Fig. 1). It also grossly overpredicts the dip.

We therefore propose a new hypothesis which places the physics of arching within a fully consistent continuum theory. We postulate that the principal axes of the stress tensor have a fixed angle of inclination τ to the downward vertical. As shown below, this 'fixed principal axes' (FPA) hypothesis corresponds to a local c.r. between stresses. At the same time, the resulting macroscopic stress predictions embody the physics of arching: the intuitive conception of lines of force arching outward from the axis is preserved, while shear forces are properly included.

Unlike other local constitutive relations^{1,2,6} the FPA closure detailed below is non-analytic on the symmetry axis $r = 0$. We believe this singularity captures an important feature of sand-pile physics: in constructing a pile (by dropping sand onto the apex), grains that have rolled down in different directions have had qualitatively different histories. More generally, we believe that the c.r. should in principle encode the local construction history of

FIG. 3 Predictions (found numerically) of the FPA model with, as the second closure equation $\sigma_{zz} = \sigma_r$ (refs 1,2,16) (solid curve); or the Haar-von Karman closure, $\sigma_{zz} = (\sigma_r + \sigma_{\theta\theta})/2 - [(\sigma_{zz} - \sigma_r)^2/4 - \sigma_{\theta\theta}^2]^{1/2}$ (dashed curve; ref. 18). Symbols are experimental points from ref. 3. Solid symbols, quartz sand; unfilled symbols, NPK-1 fertilizer, with the pile heights (h) shown. In both cases, ϕ was measured as 33° . Stresses are normalized by the total weight of the pile; there are no adjustable parameters. The prediction for f_{zz} on the central axis depends slightly on the second closure but is remarkably close to those measured for sand and for fertilizer (which themselves differ slightly). The stresses were measured on a linear array of sensors across the width of the pile; there are two data points for each symbol corresponding to two opposite points on the pile (their difference gives an indication of the experimental error)³. Note that RSF scaling is well obeyed experimentally; there is good data collapse for different pile heights. The scaling means physically that a sand pile has no characteristic length scale.



the material through its influence on the packing geometry of grains. However, unless subsequent loadings are large enough to cause slip deep within the pile, the c.r. of a given material element should be 'frozen in' at the time of its burial, after which the packing of grains within the element ceases to evolve. (From this idea, one can use RSF scaling to deduce that, for a sand-pile, the c.r. must, in cylindrical polars, be the same for all non-zero r .)

Our FPA hypothesis consists of assuming, as a constitutive relation, that the principal stress axes of a material element are fixed at the time of its burial, and unaffected by loads applied subsequently. Although ultimately its value can only be judged against experiment, this is perhaps the simplest physically reasonable model for a history-dependent local c.r.: at burial, the element remembers its construction history through a local anisotropy in grain organization, which subsequently fixes the orientation of the stresses that the element can support. As burial occurs at the surface, which is a slip plane (where the principal axes are known^{1,2}) the FPA hypothesis leads directly to the following c.r.:

$$\sigma_r = \sigma_{zz} - 2 \tan \phi |\sigma_{\theta\theta}| \quad (2)$$

The major principal axis everywhere bisects the angle between the vertical and the free surface: $\tau = (\pi - 2\phi)/4$. Although there is no inbuilt concept of arching in equation (2), on the scale of a pile, τ does define an 'arch' direction; and indeed the stresses propagate as though along a set of nested arches (Fig. 1). A detailed exposition of these findings, and a fuller discussion of all our results, including those of a more general model of which FPA is a special case, will be given elsewhere (J.P.W., P.C. and M.E.C., manuscript in preparation). As ϕ is fixed by experiment, FPA gives a complete closure of the $d = 2$ sand-pile problem. The resulting equations are easily solved. For $\sigma_{zz}(r)$, they give a piecewise linear

function with a dip at the centre (Fig. 2).

We now turn to the three-dimensional case. For a sand pile with axial symmetry, the arguments leading to equation (2) are unaffected, but a second constitutive relation is required. The simplest choice assumes a linear relation between σ_{zz} and σ_r ; as these are equal by symmetry at $r = 0$, one then has $\sigma_{zz} = \sigma_r$ everywhere^{1,2,6,16}. Results for this choice, and one alternative, are presented in Fig. 3. In fact, we have tested several secondary closures, and found that all give similar predictions. This is because the main effect of the α terms in equations (1) is to smooth out the stress profiles in the outer parts of the pile; the cusp at the centre is affected somewhat in magnitude, but not in character. (We expect the same to hold for other physically reasonable secondary closures.) Also shown in Fig. 3 are the experimental data of Smid and Novosad³. Our predictions give very good agreement with these data, including those for $r = 0$, although a deviation is seen near the maximum of σ_{zz} .

We conclude that the FPA hypothesis embodies the physically appealing picture of arches^{4,5,14}, while providing a powerful new closure scheme for the sand pile. Considering the absence of adjustable parameters, the fit with experiment is remarkably good. More generally, the FPA model exemplifies a new modelling strategy for static granular media: we assume the existence of local constitutive relations among stresses, but acknowledge their dependence on the construction history of the pile. Although equation (2) is specific to the sand-pile geometry, the same basic approach may be applicable to many other problems in granular mechanics.

In future we aim to include the effects of randomness (noise) in our equations. This has been studied recently using discrete models (ref. 17, and P. G. de Gennes, personal communication) whose relation to this work will be discussed elsewhere. □

Received 22 December 1995; accepted 11 June 1996.

- Nedderman, R. M. *Statics and Kinematics of Granular Materials* (Cambridge Univ. Press, 1992).
- Sokolovskii, V. V. *Statics of Granular Materials* (Pergamon, Oxford, 1965).
- Smid, J. & Novosad, J. in *Proc. of 1981 Powtech. Conf., Ind. Chem. Eng. Symp.* **63**, D3V 1-12 (1981).
- Edwards, S. F. & Oakeshott, R. B. *Physica D* **38**, 88-93 (1989).
- Edwards, S. F. & Mounfield, C. C. *Physica A* **226**, 1-33 (1996).
- Bouchaud, J.-P., Cates, M. E. & Claudin, P. *J. Physique* **5**, 639-656 (1995).
- Bouchaud, J.-P., Cates, M. E., Prakash, J. R. & Edwards, S. F. *Phys. Rev. Lett.* **74**, 1982-1985 (1995).
- Brown, R. L. & Richard, J. C. *Principles of Powder Mechanics* (Pergamon, New York, 1966).
- Jokati, T. & Moriyama, R. J. *Soc. Powder Technol. Jpn* **60**, 184-189 (1979).
- Bagster, D. F. & Kirk, R. J. *Powder Bulk Solids Tech.* **1**, 19-24 (1985).
- Hong, D. C. *Phys. Rev. E* **47**, 760-762 (1993).

- Huntley, J. *Phys. Rev. E* **48**, 4099-4111 (1993).
- Liffman, K., Chan, D. Y. & Hughes, B. D. *Powder Technol.* **78**, 263-271 (1994).
- Trolope, D. H. in *Rock Mechanics in Engineering Practice* (eds Stagg, K. G. & Zienkiewicz, O. C.) 275-320 (Wiley, New York, 1968).
- Trolope, D. H. & Burman, B. C. *Geotechnique* **30**, 137-157 (1980).
- Claudin, P. thesis, Univ. Cambridge (1989).
- C. H. Liu et al. *Science* **269**, 513-515 (1995).
- Wood, D. M. *Soil Behaviour and Critical State Soil Mechanics* (Cambridge Univ. Press, 1990).

ACKNOWLEDGEMENTS. This work was funded in part by EPSRC and by the Colloid Technology Programme. M.E.C. acknowledges the hospitality of the Newton Institute, Cambridge, where part of this work was done.

CORRESPONDENCE should be addressed to M.E.C. (e-mail: m.cates@ed.ac.uk).

Using antibodies to perturb the coordination sphere of a transition metal complex

Phalguni Ghosh*, Doron Shabat†, Sandeep Kumar†, Subhash C. Sinha†, Flavio Grynszpan*, Jian Li*, Louis Noodleman* & Ehud Keinan*†

* Department of Molecular Biology, The Scripps Research Institute, 10666 N. Torrey Pines Road, La Jolla, California 92037, USA

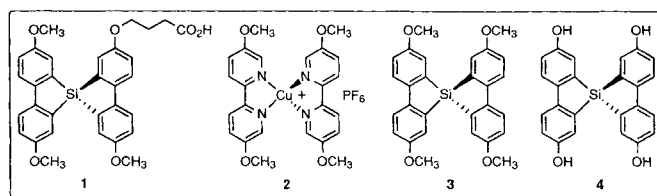
† Department of Chemistry, Technion—Israel Institute of Technology, Technion City, Haifa 32000, Israel

METAL ions in the active sites of many metalloenzymes exhibit distinctive spectral and chemical features which are different from those of small inorganic complexes^{1,2}. These features are the result of the unusual geometric and electronic constraints that are imposed on the metal ion within the protein environment³. Much effort has been invested to try to mimic this feature of metalloenzymes in synthetic systems, but this remains a formidable task. Here we show that one of the key lessons learned from the science of catalytic antibodies—that binding energy can be converted into chemical energy⁴—can be exploited to ‘fine-tune’ the physicochemical properties of a metal complex. We show that an antibody’s binding site can reversibly perturb the coordination geometry of a metal ion, and can stabilize a high-energy coordinated species⁵. Specifically, antibodies designed to bind the organosilicon compound **1** (Fig. 1) also bind the geometrically similar Cu(I) complex **2**. However, the antibody binds a slightly compressed form of **2**, which is closer in size to **1**. This distortion is manifested by a spectral shift—an ‘immunochromic’ effect.

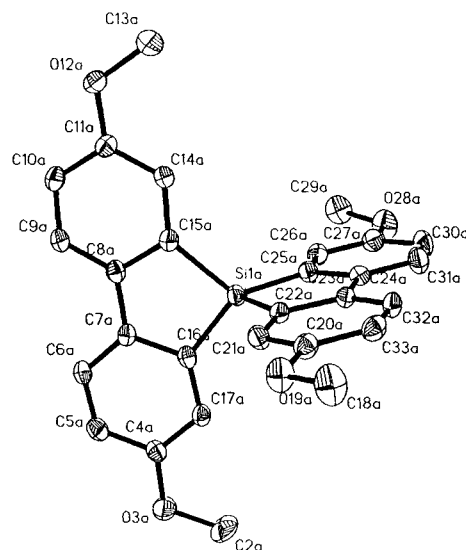
Spirosilanes **1**, **3** and **4** and the copper (I) complex **2**, as well as the protein conjugates of **1**, were synthesized as described (see Supplementary Information). The crystal structure of **3** (Fig. 1) shows that the two biphenyl planes are orthogonal (dihedral angle, 90.2°). In contrast, the corresponding two bipyridine planes in the crystal structure of **2** are not (dihedral angle, 74°). As tetracoordinated (bipy)₂Cu(I) complexes are known to be tetrahedral in solution^{6,7} the observed deviation from the tetrahedral geometry in the solid state is probably caused by crystal packing forces. Whereas the average Si–C bond length in **3** is about 1.87 Å, the corresponding Cu–N bond length in **2** is ~2.03 Å. Thus, compounds **2** and **3** have similar molecular shapes in solution, differing mainly in the overall size of the molecule.

Following immunization of 129 G^{IX+} mice with the ovalbumin conjugate-**1**, five hybridoma cells producing anti-**1** antibodies were selected for further studies on the basis of screening by enzyme-linked immunosorbent assay (ELISA) for affinity to the bovine serum albumin (BSA)-**1** conjugate. Antibodies from each cell line were purified by ammonium sulphate precipitation, anion exchange, and protein-G affinity chromatography⁸.

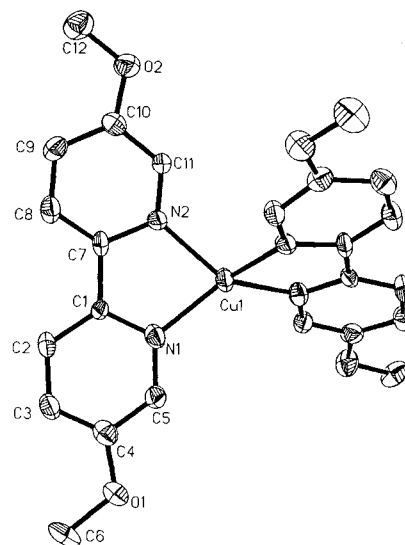
The ultraviolet–visible spectrum of **2** in a phosphate buffer (pH 7.4) is characterized by a distinct, metal-to-ligand charge-transfer band at 405 nm^{9,10}. Upon addition of stoichiometric amounts of antibody 8A12, this band is red-shifted to 460 nm (Fig. 2a). As a highly hydrophobic hapten such as **1** is expected to induce a hydrophobic combining site in the antibody¹¹, a simplistic explanation of the observed spectral changes could be based on medium effects^{12,13}. We have already reported on similar shifts (up to 8 nm) in the Soret band of metalloporphyrins upon binding to their specific antibodies¹⁴. Small redshifts (up to 12 nm) are observed when the medium is changed from water to organic solvents (in the order solvent, peak wavelength $\lambda_{\max}$ (nm), extinction coefficient ϵ : water at pH 7.4, 405, 1,400; acetone, 406, 1,400; 4-trifluoromethylchlorobenzene, 411, 1,400; methanol, 413, 3,000; dichloromethane, 414, 2,000; dimethylformamide (DMF),



a



b



c

	2	3
Space group	C2	P1̄
a	27.606 (9) Å	16.965 (8) Å
b	6.950 (4) Å	14.768 (7) Å
c	15.690 (7) Å	11.196 (6) Å
α	90°	105.75° (5)
β	92.48° (3)	76.46° (5)
γ	90°	104.90° (5)
Z	4	4
R factor	0.078	0.087

FIG. 1 a, X-ray crystallographic structure and atom numbering of spirosilane **3**. b, Crystallographic structure and atom numbering of Cu(I) complex **2**. c, Selected crystallographic parameters of **2** and **3**; see Supplementary Information for full crystallographic data.

417, 710). Consequently, only a small portion of the 55 nm redshift observed upon binding of **2** to antibody 8A12 can be explained by medium effects alone.

We propose an alternative explanation to the observed redshift on the basis of the structural differences between the spiroilane hapten and copper complex **2**. Although both compounds share the same general shape, the size of compound **2** is larger than that of **1**. Hence, the binding site of an antibody to **1** may be too small to accommodate the copper complex **2**. Yet such an antibody may bind **2** if part of the binding energy is used to stabilize a compressed, high-energy state of **2**.

We have carried out quantum chemical calculations on the electronic spectra^{15–18} of two model systems in order to assess whether shortening of all four metal–ligand bonds by 0.16 Å could explain the observed redshift. The dipole-allowed transition bands were calculated from the energy difference between ground state and metal-to-ligand charge transfer (MLCT) excited singlet state by a non-local density functional method. Vertical excitation energies were calculated in accordance with the expected position of the band maximum based on the Franck–Condon principle. For the simplified model bis(diiminoethane)copper(I), **5**, the calculated geometry-optimized Cu–N bond length of the ground

state is 2.07 Å. Upon shortening of these bonds to 2.01 Å, the band at 422 nm, which corresponds to transition ($d_{xz}: b_1$) $\rightarrow$ ($\pi^*: b_2$ or b_3) 1B_3 or 1B_2 (in the point group D_2), is slightly shifted to 431 nm (Fig. 2b). Further compression of this complex down to a bond length of 1.88 Å increases the redshift of that specific band to 467 nm. Considering the fact that a redshift of ~10 nm is expected because of hydrophobic medium effects, this 45 nm shift agrees quite well with the observed immunochromic effect of 55 nm. Similar calculations of this band were done for another complex bis(5,5'-dihydroxy-2,2'-bipyridine)copper(I), **6**, which is a more realistic model for **2**. A similar redshift was obtained as well for this specific band, but the magnitude of the shift was smaller, from 400 nm at 2.01 Å of Cu–N bond length to 415 nm at 1.88 Å.

The calculated redshift in the ($d_{xz}: b_1$) $\rightarrow$ ($\pi^*: b_2$ or b_3) transition energy for both complexes arises from destabilizing the ground state more than the 1B_3 excited state as the complexes are compressed. The spectral shifts arise from the change in the energy difference between ground and excited states, but the destabilization of the ground state alone is particularly interesting as this gives the strain energy on the Cu(I) complex when it binds to the antibody. From our calculations, the energy cost of compressing the Cu–N bond length from 2.07 Å (optimized) in **5** to 1.88 Å is about 15 kcal mol^{–1}. Similarly, changing Cu–N from 2.01 Å to 1.88 Å in **6** costs ~12 kcal mol^{–1}. By comparison, the spectral shifts for the same bond length changes are 6.7 and 2.6 kcal mol^{–1} for complexes **5** and **6**, respectively, showing that a large part of the spectral shift arises mainly from ground-state destabilization. The strain energy for compressing the Cu(I) complex in that antibody binding site probably occurs at the expense of the Cu(I) complex–antibody binding energy. As indi-

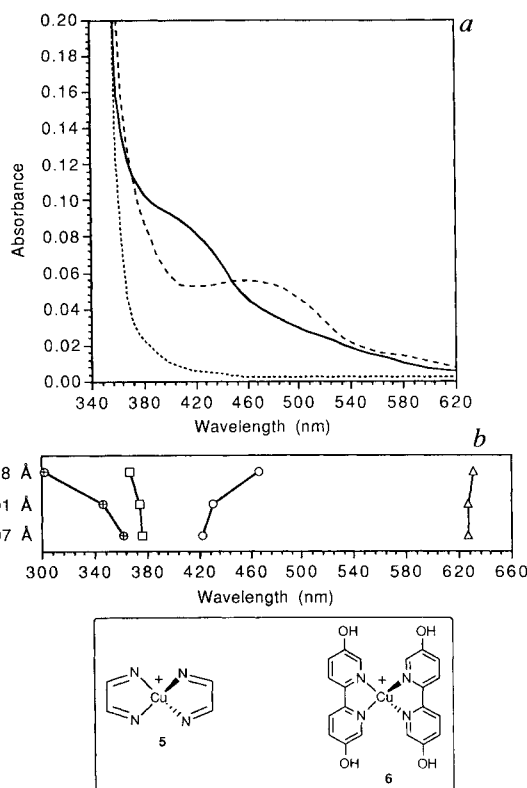
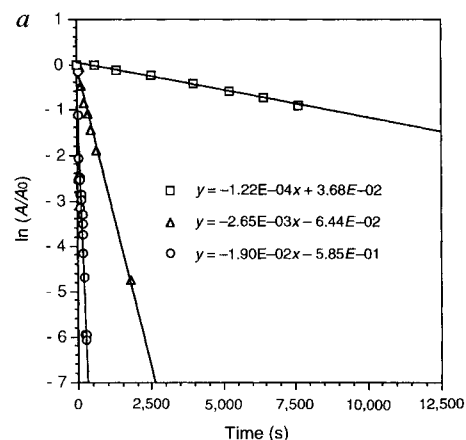


FIG. 2 a, UV-vis spectra of the copper complex. Solid line, absorption of **2**; dashed line, absorption of **2** in the presence of excess antibody 8A12; dotted line, absorption of **2** after exposure to air. b, Calculated ultraviolet-visible spectra of copper(I) complex **5** at various Cu–N bond lengths. Calculations were carried out using the Amsterdam Density Functional (ADF) package (Department of Theoretical Chemistry, Vrije Universiteit, Amsterdam). The local density approximation for exchange and correlation used the parametrization of Vosko *et al.*²⁵. The non-local corrections, which are based on Becke's gradient correction to exchange²⁶ and Perdew's²⁷ correction to correlation, were added in each self-consistent cycle. The numerical integration scheme adopted was the polyhedron method²⁸. A set of uncontracted triple- ζ Slater-type orbitals (STO) plus polarization function (ADF basis set IV) was employed for the valence orbitals. The inner core shells were treated by the frozen core approximation. A set of auxiliary *s*, *p*, *d*, *f* and *g* STO functions was introduced to fit the molecular density and to represent Coulomb and exchange potentials accurately. The energies of singlet excited states were calculated by the method of Zeigler *et al.*²⁹. All calculations were done with a spin-unrestricted scheme.



b

Antibody	K_d of 1 *	K_d of 4 *	Air-ox. rate†	Redshift‡	CV shift§
5E10	1.6×10^{-7}	9.5×10^{-7}	2×10^{-4}	20	140
8D7	1.6×10^{-6}	3.7×10^{-6}	6×10^{-5}	22	75
8A12	1.3×10^{-6}	4.8×10^{-6}	1×10^{-4}	55	80
9C9	2.2×10^{-7}	3.2×10^{-6}	8×10^{-5}	24	ND
5A1	5.7×10^{-7}	2.4×10^{-6}	2×10^{-4}	22	ND

FIG. 3 a, Determination of pseudo-first-order rate constants of air oxidation of **2**. All experiments were carried out in phosphate buffer (50 mM, pH 7.4). The Cu(I) complex was prepared *in situ* either using complex **2** or by reduction of the corresponding Cu(II) complex, with slight excess of sodium dithionate under argon. The concentration of the Cu(I) species was monitored by its absorption at 400 nm after exposing the solution to air. □, Oxygenation of **2** (0.05 mM) in the presence of 8A12 (0.07 mM); △, oxygenation of **2** in the presence of 8A12 and **1**; ○, oxygenation of **2** alone in phosphate buffer solution. b, * Dissociation constants (M); † pseudo-first-order rate constants (s^{–1} for Cu(I) oxidation); ‡ shift of the ultraviolet-visible absorption (nm); § shift of the anodic peak potential (mV) measured by cyclic voltammetry (CV); ND, not determined.

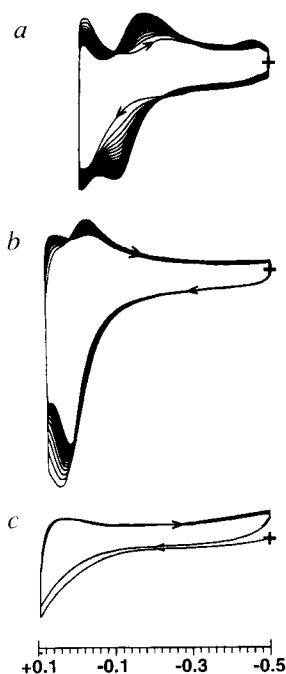


FIG. 4 Cyclic voltammograms with a hanging mercury drop electrode versus Ag/AgCl, in phosphate buffer, pH 7.4, at a scan rate of 100 mV s^{-1} within a potential range of between -0.5 and $+0.1 \text{ V}$. *a*, Cyclic voltammogram of **2**; *b*, cyclic voltammogram of **2** in the presence of antibody 5E10; *c*, cyclic voltammogram of the solution in *b* 30 s after addition of **1** (3 equivalents).

cated above, at a dihedral angle of 90° between the two bipyridine planes, the transition at 422 nm is doubly degenerate (in fact, D_2 is a subgroup of S_4). However, this and other bands split at angles different from 90° . This splitting is expected to produce a slight spectral broadening (of the order of $4\text{--}8 \text{ nm}$) of the observed bands when the dihedral angle is smaller than 90° (for example, 74° in the crystallographic structure of **2**). Most importantly, changes in the dihedral angle do not produce a significant spectral shift in the relevant band.

The affinity (dissociation constants) of **1** and **4** to all five antibodies were determined by the fluorescence quenching method¹⁹. As we expect (Fig. 3*b*), the affinity of these antibodies to spiroisilane **1** are uniformly higher than those measured with the slightly smaller spiroisilane **4**.

The air oxidation of **2** in water solution is very fast (at pH 7.4, the pseudo-first-order rate constant is $1.9 \times 10^{-2} \text{ s}^{-1}$). This one-electron oxidation of Cu(I) to Cu(II) involves major changes in the coordination geometry. Whereas tetracoordinated Cu(I) complexes are tetrahedral, the corresponding Cu(II) complexes are quasi-square-planar in the ground state^{20,21}. Oxidation does not involve a significant change in the copper–ligand bond length (shortening by $0.02\text{--}0.04 \text{ \AA}$)^{22,23}. We found that this oxidation reaction is significantly inhibited by the presence of any of the above-mentioned five antibodies (Fig. 3). Control experiments show that in the presence of haptin **1**, complex **2** is no longer 'protected' by the antibody against air oxidation. These observations confirm that **2** is specifically bound at the antibody binding site. It seems likely that the air-oxidation rate of **2** is related to its affinity constant, the slower the oxidation-rate, the stronger the binding to the antibody.

Whereas oxygen seems to be an oxidant too weak to oxidize the Cu(I) complex **2** while inside the antibody binding site, this task could be achieved by electrochemical oxidation. The electrochemical behaviour of **2** (Fig. 4) was studied by cyclic voltammetry using a hanging mercury drop electrode versus Ag/AgCl electrode in chloride-free phosphate buffer (50 mM , pH 7.4) under an argon atmosphere. In the absence of any antibody, complex **2** exhibits a single, reversible redox couple ($E_{1/2} = -135 \text{ mV}$; $\Delta E_p = 60 \text{ mV}$).

The peak currents increase gradually and reach a maximum within 3 min, suggesting that the redox process involves absorption of the electroactive species onto the mercury surface²⁴.

The cyclic voltammogram of **2** changes remarkably upon addition of an anti-**1** antibody. Thus, in the presence of stoichiometric amounts of antibodies 5E10, 8D7 and 8A12, the formal potentials of **2** are shifted to more positive potentials (Fig. 4*b*). These shifts reflect preferential stabilization of the Cu(I) over the Cu(II) complex by the three antibodies. Interestingly, the observed anodic oxidation peak is much more intense than the cathodic reduction peak in the reverse cycle. This behaviour suggests that the initial oxidation product is a tetrahedral Cu(II) complex. This high-energy species is poorly bound to the antibody and dissociates within less than 5 seconds. Addition of haptin **1** to a mixture containing both **2** and antibody 5E10 causes complete inhibition of the electron-transfer process (Fig. 4*c*). This can be explained by insulation of the electrode by a layer of neutralized antibody molecules, as was confirmed by control experiments using non-relevant proteins such as BSA.

Metal-containing organic and inorganic complex catalysts are fascinating synthetic tools because their reactivity can be adjusted as a function of coordination number, coordination geometry, oxidation state, symmetry, chirality, and so on. Chemists apply two design elements to control these parameters in organometallic catalytic complexes: selection of an appropriate metal nucleus, and selection and synthetic modifications of ligands with appropriate stereoelectronic properties. In metalloenzymes, nature uses a third control element, probably a more effective one, to fine-tune the desired metal reactivity: building a supramolecular envelope around the metal centre. The design and synthesis of such envelopes seems to be still beyond the reach of chemists. We have shown here that antibodies can be induced to perturb the coordination sphere of a given metallic complex and thereby modify its physicochemical properties. Future work along these lines may lead us to a whole new class of metalloproteins. □

Received 22 March; accepted 4 June 1996.

1. Vallee, B. L. & Williams, R. J. P. *Proc. natn. Acad. Sci. U.S.A.* **59**, 498–505 (1968).
2. Williams, R. J. P. *Eur. J. Biochem.* **234**, 363–381 (1995).
3. Solomon, E. I. & Lowery, M. D. *Science* **259**, 1575–1581 (1993).
4. Schultz, P. G. & Lerner, R. A. *Science* **269**, 1835–1842 (1995).
5. Shokat, K. M., Leumann, C. J., Sugawara, R. & Schultz, P. G. *Angew. Chem. int. Edn engl.* **27**, 1172–1174 (1988).
6. Nabeshima, T., Inaba, T., Furukawa, N., Hosoya, T. & Yano, Y. *Inorg. Chem.* **32**, 1407–1416 (1993).
7. Kitagawa, S. & Munakata, M. *Inorg. Chem.* **20**, 2261–2267 (1981).
8. Kohler, G. & Milstein, C. *Nature* **256**, 495–497 (1975).
9. Williams, R. J. P. *J. chem. Soc.* 137–145 (1955).
10. Palmer, R. A. & Piper, T. S. *Inorg. Chem.* **5**, 864–878 (1966).
11. Shaab, D., Itzhaky, H., Raymond, J.-L. & Keinan, E. *Nature* **374**, 143–146 (1995).
12. Lewis, C. T., Krämer, T., Robinson, S. & Hilvert, D. *Science* **253**, 1019–1022 (1991).
13. Lewis, C. T., Paneth, P., O'Leary, M. H. & Hilvert, D. *J. Am. chem. Soc.* **115**, 1410–1413 (1993).
14. Keinan, E. et al. *Inorg. Chem.* **31**, 5433–5438 (1992).
15. Zeigler, T. *Can. J. Chem.* **73**, 743–761 (1995).
16. Rosa, A. & Baerends, E. J. *Inorg. Chem.* **33**, 584–595 (1994).
17. Noodleman, L. & Baerends, E. J. *J. Am. chem. Soc.* **106**, 2316–2327 (1984).
18. Jones, D. H., Hinman, A. S. & Ziegler, T. *Inorg. Chem.* **32**, 2092–2095 (1993).
19. Steward, M. W. & Steensgaard, J. in *Antibody Affinity: Thermodynamic Aspects and Biological Significance* 76–77 (CRC Press, Boca Raton, Florida, 1983).
20. Karlin, K. D. & Yandell, J. K. *Inorg. Chem.* **23**, 1184–1188 (1984).
21. Hathaway, B. J. in *Comprehensive Coordination Chemistry* Vol. 5 (ed. Wilkinson, G.) 533–774 (Pergamon, New York, 1987).
22. Johnson, J. E., Beineke, T. A. & Jacobson, R. A. *J. chem. Soc. (A)* 1371–1374 (1971).
23. Foley, J., Tyagi, S. & Hathaway, B. J. *J. chem. Soc. Dalton Trans.* 1, 1–5 (1994).
24. Bard, A. J. (ed.) in *Encyclopedia of Electrochemistry of the Elements* IXA 172–211 (Dekker, New York, 1983).
25. Vosko, S. H., Wilk, L. & Nusair, M. *Can. J. Phys.* **58**, 1200–1211 (1980).
26. Becke, A. D. *J. chem. Phys.* **84**, 4524–4529 (1986).
27. Perdew, J. P. *Phys. Rev. B* **33**, 8822–8824 (1986).
28. te Velde, G. & Baerends, E. J. *J. comput. Phys.* **99**, 84–98 (1992).
29. Ziegler, T., Rauk, A. & Baerends, E. J. *Theor. chim. Acta* **43**, 261–273 (1977).

ACKNOWLEDGEMENTS. We thank the US NIH for grants to E.K. and L.N., and the US–Israel Binational Science Foundation for financial support. F.G. is a Fellow of the Lady Davis Foundation Fellowship Trust. We also thank D. Kubitz at the antibodies, F. C. Anson for his help with the electrochemical experiments, and R. Chadha and M. Kapon for the X-ray crystal structure analysis.

SUPPLEMENTARY INFORMATION is available on Nature's World-Wide Web site <http://www.nature.com>. Paper copies are available from Mary Sheehan at the London Nature (editorial) office.

CORRESPONDENCE and requests for materials should be addressed to E.K. (e-mail: keinan@scripps.edu).

Sulphur isotope fractionation in modern microbial mats and the evolution of the sulphur cycle

Kirsten S. Habicht & Donald E. Canfield

Max Planck Institute for Marine Microbiology, Celsiusstrasse 1, 28359 Bremen, Germany

THE sulphur cycle has evolved over the course of the Earth's history^{1,2}. The early Earth's surface environment was reducing, containing little atmospheric oxygen³, and with seawater sulphate concentrations estimated at less than a few per cent of those found today. The accumulation of sulphate in the ocean to much higher concentrations was probably coincident with the initial accumulation of oxygen in the atmosphere and the consequent oxidative weathering of sulphide minerals on land^{4,5}. Past changes in sulphate concentrations in ancient oceans have previously been assessed by comparing the systematics of sulphur isotope fractionation by sulphate-reducing bacteria⁶⁻⁹ with the isotopic composition of sedimentary sulphides^{1,2,5,10,11}. But such interpretations have proven equivocal: the generally small ³⁴S depletions in Archaean sulphides (deposited ~2.5–3.8 billion years ago) have been separately argued to result both from rapid sulphate reduction in a sulphate-rich ocean^{5,12}, and from sulphide formation in a sulphate-poor ocean^{1,2,11}. Here we report large ³⁴S depletions of 20–25%, observed during rapid sulphate reduction by sulphate-reducing bacteria in modern photosynthetic cyanobacterial mats from Solar Lake, Sinai. We conclude that high sulphate concentrations give rise to highly ³⁴S-depleted sulphides, and thus that appreciable concentrations of seawater sulphate did not accumulate until the initial accumulation of oxygen into the atmosphere in post-Archaean times.

Sulphate-reducing bacteria catalyse the oxidation of organic carbon using sulphate as an electron acceptor, producing sulphide isotopically depleted in the heavy isotope ³⁴S (refs 6–8). With sulphate levels >1 mM, pure cultures of sulphate-reducing bacteria produce sulphides depleted in ³⁴S by 4–46‰ (refs 6–8,13), with an average of 18‰ (ref. 14). Only limited ³⁴S depletion (<4‰) is observed, however, when sulphate reaches low levels <1.0 mM (ref. 15). Sedimentary sulphides from the Archaean and lower Proterozoic have relatively consistent $\delta^{34}\text{S}$ values of

$\sim 0 \pm 5\text{‰}$, similar to meteorites and mantle-derived igneous rocks^{2,11,16–20}, and that for contemporaneous sea water^{1,16}. These sulphides have therefore been attributed to either bacterial sulphate reduction in low concentrations of seawater sulphate or to a mantle origin^{2,17–19,21}.

Alternatively, with abundant sulphate, pure cultures of sulphate-reducing bacteria produce the most ³⁴S-depleted sulphides at low specific rates of sulphate reduction (rate per unit cell), with ³⁴S depletion decreasing as the rate increases^{6,8}. Also, a tendency has been noted for the depletion of modern sedimentary sulphides in ³⁴S to decrease as the sulphate-reduction rate increases (rate per volume)^{5,9} (but see refs 14,22). These observations have been extrapolated, and used to argue that the small isotopic differences between sedimentary sulphides and coeval sulphate in Archaean oceans have resulted from small ³⁴S depletions during rapid sulphate reduction in a sulphate-rich ocean^{5,12}. Sulphate reduction rates of between 27 to 270 $\mu\text{mol cm}^{-3} \text{d}^{-1}$ have been suggested⁵. Thus, two very different scenarios for ocean chemistry have been generated from essentially the same isotopic results.

The choice between these two scenarios, and subsequently high or low levels of sulphate in Archaean oceans, rests on the ability of natural populations of sulphate-reducing bacteria to fractionate at high rates of sulphate reduction, for which no determinations have previously been reported. To provide this information, we measured the fractionation of sulphur isotopes during sulphate reduction in natural populations of sulphate-reducing bacteria from the microbial mats of Solar Lake, Sinai. Our choice of a mat environment is driven by two considerations. First, modern microbial mats are viewed as reasonable homologues to ancient benthic, photosynthetic, microbial communities, including stromatolites²³. Second, microbial mats support the highest rates of sulphate reduction (per volume) yet reported on Earth^{24,25}. Extreme rates of sulphate reduction are driven by high rates of phototrophic carbon fixation, which, when depth-integrated, are similar to those measured in coastal marine waters²⁶, but are compressed into a zone extending over only 0.5–3 mm depth^{25,27}. Most of the photosynthetically fixed carbon in cyanobacterial mats, like those of Solar Lake, is oxidized by sulphate-reducing bacteria^{25,27}. Thus, the direct transfer of large amounts of fresh photosynthetically produced organic carbon to sulphate reducers, over small vertical dimensions, is responsible for the high rates of sulphate reduction. An even more efficient transfer of carbon from phototrophs to heterotrophic carbon oxidizers in ancient microbial mats and stromatolites, providing higher rates of sulphate reduction, is not anticipated. Hence, the high rates of sulphate reduction

FIG. 1 Depletion of sulphide in ³⁴S, compared to sulphate, versus sulphate-reduction rate (SRR) for natural populations of sulphate-reducing bacteria from the microbial mats of Solar Lake, Sinai, at different temperatures and sulphate concentrations. The top 0–13 mm of the mats were sectioned into 5 depth intervals. Each mat piece was placed into a dialysis membrane and incubated in a gas-tight bag filled with N₂-purged Solar Lake water (65 mM sulphate) or diluted Solar Lake water (20 mM sulphate) and stored in the dark at 10 °C, 20 °C or 30 °C. Sulphate-reduction rates were calculated from sulphide accumulated in the bag²³. Rates were generally linear with time, except in the surface interval where they accelerated during incubation. The accelerated rates could have resulted from the accelerated decay of fresh cyanobacteria. After 1.5 to 8 days, the bag-water was withdrawn, $\delta^{34}\text{S}$ values of the sulphate in the bag and the sulphide produced were measured, and calculated relative to Canyon Diablo Troilite (CDT). In all cases, less than 1% of the initial sulphate was reduced during incubation, ensuring open-system conditions for determining ³⁴S sulphide depletion. The *in situ* rates of sulphate reduction were measured using ³⁵S-sulphate (ref. 24). Our bag rates at 20 °C were comparable to the *in situ* rates, except in surface depth interval, where the bag rates were higher. In addition to the possible enhanced decomposition of fresh cyanobacteria in our bag incubations, *in situ* rates in the surface depth interval may have been underestimated owing to complications of rapid sulphide oxidation during the rate determination²⁴.

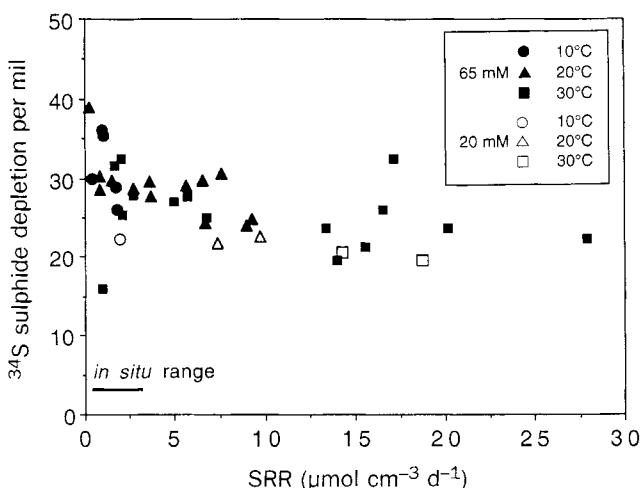


TABLE 1 Isotope compositions of sulphides from Archaean stromatolites and Solar Lake microbial mat

Sample	Age (Gyr)	$\delta^{34}\text{S}\text{‰}$	Location
PPRG 1408	3.0	1.5 (2)	Insuzi group, Natal, South Africa
PPRG 2119	2.75	-1.5 (8.5)	Tumbiana formation, Hamersley basin, Western Australia
PPRG 2079	2.75	0.5 (6.5)	Tumbiana formation, Hamersley basin, Western Australia
Microbial mat	Modern	-16 to -22 (-36 to -44)	Solar Lake, Sinai

Pyrite sulphur from Archaean stromatolites was liberated for isotope analysis by chromium reduction³¹ of kerogen extracts. Range in isotope values for Solar Lake mat is from both acid-volatile sulphide phases (mostly FeS) and pyrites. Depletions of sulphides in ^{34}S compared to contemporaneous seawater sulphate are shown in parentheses, with estimates for sulphate of 3.5‰ at 3.3 Gyr (ref. 18), and 7‰ at 2.7 Gyr (ref. 32). Isotope composition of Solar Lake sulphate is 20‰.

measured in modern microbial mats should range among the highest rates possible in ancient sedimentary environments.

In the cyanobacterial mats of Solar Lake we measured *in situ* sulphate reduction rates of $0.3\text{--}3.1\ \mu\text{mol cm}^{-3}\text{ d}^{-1}$, at an ambient temperature of 20°C (Fig. 1), consistent with previous rate determinations on these mats²⁸. To encounter a broader range in sulphate-reduction rates, we incubated 2–3-mm slices from the surface 1.3 cm of the mat at temperatures from 10 to 30°C . We also explored the effect of variable sulphate concentration (65 mM, the concentration of Solar Lake water, and 20 mM) on both rates and fractionation. No difference in sulphate-reduction rate was encountered when the same mat piece was incubated at both 65 mM and 20 mM sulphate ($n = 5$). Overall, rates of sulphate reduction in our experiments ranged from 0.2 to $28\ \mu\text{mol cm}^{-3}\text{ d}^{-1}$. The high rates are among the highest ever reported²⁵ and are also within the lower range proposed for the Archaean ocean⁵ (see above).

The depletion of sulphide in ^{34}S varied from 16 to 39‰ (Fig. 1). These values compare with those obtained from pure cultures of sulphate-reducing bacteria^{6–8,13}. At a sulphate concentration of 65 mM, the highest ^{34}S depletions were generally associated with low rates of sulphate reduction, also consistent with pure culture studies^{6,8,13}. This correlation, however, was not observed at 20 mM

sulphate, where relatively consistent ^{34}S depletions were found through a broad range of sulphate-reduction rates. For both sulphate concentrations, sulphide was depleted by 20–32‰ at high rates of $\geq 10\ \mu\text{mol cm}^{-3}\text{ d}^{-1}$. These observations indicate that sulphide is produced in the mat with a minimum ^{34}S depletion of about 20‰, independent of sulphate-reduction rate, and sulphate concentration in the range of 20 to 65 mM. This value is greater than the small (0–5‰) ^{34}S depletions inferred for high rates of sulphate reduction in a sulphate-rich Archaean ocean^{5,12}.

Further, large ^{34}S depletions of 36 to 42‰ are measured for solid-phase sulphides in Solar Lake mats (Table 1). These sulphides are fixed in the surface few centimetres of the mat under open conditions with minimal sulphate depletion. Thus, the influence of isotope fractionation by sulphate-reducing bacteria is preserved, with perhaps some small additional fractionation induced by the oxidative sulphur cycle^{14,22}. Isotope analysis of pyrites from Archaean stromatolites from the PPRG (Precambrian Paleobiology Research Group) sample collection show values of between -1.5 to 1.5‰. These values indicate ^{34}S depletions, compared to best-estimates for contemporaneous seawater sulphate, of between 2 to 8.5‰ (Table 1). Such low ^{34}S depletions are far less than those measured during sulphate reduction in the Solar Lake mats (Fig. 1), or those for sulphides preserved in the mats (Table 1).

In summary, our results demonstrate that sulphide is depleted in ^{34}S by $>20\text{‰}$ at high rates of sulphate reduction in natural populations of sulphate-reducing bacteria exposed to sulphate levels $>20\text{ mM}$. Thus, the generally small isotopic differences between sulphides from Archaean and early Proterozoic sediments and stromatolites, and contemporaneous seawater sulphate, cannot be explained by sulphide formation at high rates of sulphate reduction in a sulphate-rich ocean. Instead, as previously proposed^{2,21}, sedimentary sulphides form this time originated from either biological or mantle sources, in an ocean low in sulphate. The increase in seawater sulphate concentration is, thus, best represented by the large negative isotope shift in shale-hosted sedimentary sulphides at between 2.2 to 2.3 Gyr (refs 2,17–19,21). This timing correlates with the initiation of a large early Proterozoic burial pulse of organic carbon. This burial pulse released oxidants to the Earth surface²⁹, first as oxygen, and subsequently as sulphate and Fe(III), as the oxygen reacted with reduced species on land and dissolved in the ocean³⁰. Thus, the history of atmospheric oxygen and seawater sulphate are intimately linked^{4,5,17}. This early oxygenation, however, represented only the first stage in the build-up of oxygen to present-day levels. Oxygen levels of greater than 5 to 18% of present levels were probably not attained until ~ 0.8 Gyr, near to when Metazoan life is presumed to have evolved¹⁴. □

Received 5 February; accepted 12 June 1996.

- Schidlowski, M., Hayes, J. M. & Kaplan, I. R. in *Earth's Earliest Biosphere: Its Origin and Evolution* (ed. Schopf, J. W.) 149–186 (Princeton Univ. Press, Princeton, New Jersey, 1983).
- Cameron, E. M. *Nature* **296**, 145–148 (1982).
- Kasting, J. F. *Science* **259**, 920–926 (1993).
- Walker, J. C. G. & Brimblecombe, P. *Precamb. Res.* **28**, 205–222 (1985).
- Ohmoto, H., Kakegawa, T. & Lowe, D. R. *Science* **262**, 555–557 (1993).
- Kaplan, I. R. & Rittenberg, S. C. *J. gen. Microbiol.* **34**, 195–212 (1964).
- Kemp, A. L. W. & Thode, H. G. *Geochim. cosmochim. Acta* **32**, 71–91 (1968).
- Chambers, L. A., Trudinger, P. A., Smith, J. W. & Burns, M. S. *Can. J. Microbiol.* **21**, 1602–1607 (1975).
- Goldhaber, M. B. & Kaplan, I. R. *Soil Sci.* **119**, 42–55 (1975).
- Goodwin, A. M., Monster, J. & Thode, H. G. *Econ. Geol.* **71**, 870–891 (1976).
- Hattori, K., Krouse, H. R. & Campbell, F. A. *Science* **221**, 549–551 (1983).
- Ohmoto, H. & Felder, R. P. *Nature* **328**, 244–246 (1987).
- Harrison, A. G. & Thode, H. G. *Trans. Faraday Soc.* **54**, 84–92 (1958).
- Canfield, D. E. & Teske, A. *Nature* **382**, 127–132 (1996).
- Harrison, A. G. & Thode, H. G. *Trans. Faraday Soc.* **53**, 1648–1651 (1957).
- Schidlowski, M. *Origins of Life* **9**, 299–311 (1979).
- Hattori, K., Campbell, F. A. & Krouse, H. R. *Nature* **302**, 323–326 (1983).
- Lambert, I. B. & Donnelly, T. H. in *Stable Isotopes and Fluid Processes in Mineralization* (eds Herbert, H. K. & Ho, S. E.) 260–268 (Geology Department and University Extension, Perth, Australia, 1990).
- Monster, J. et al. *Geochim. cosmochim. Acta* **43**, 405–413 (1979).

- Donnelly, T. H. et al. *J. geol. Soc. Australia* **24**, 409–420 (1977).
- Hayes, J. M., Lambert, I. B. & Strauss, H. in *The Proterozoic Biosphere. A Multidisciplinary Study* (eds Schopf, J. W. & Klein, C.) 129–132 (Cambridge Univ. Press, Cambridge, 1992).
- Canfield, D. E. & Thamdrup, B. *Science* **266**, 1973–1975 (1994).
- Des Marais, D. J. *Trends Ecol. Evol.* **5**, 140–143 (1990).
- Jørgensen, B. B. *FEMS Microbiol. Ecol.* **13**, 303–312 (1994).
- Canfield, D. E. & Des Marais, D. J. *Geochim. cosmochim. Acta* **57**, 3971–3984 (1993).
- Harrison, A. W. et al. in *Coastal Upwelling* (ed. Richards, F. A.) 303–311 (American Geophysical Union, Washington DC, 1981).
- Revsbech, N. P., Jørgensen, B. B. & Blackburn, T. H. *Limnol. Oceanogr.* **28**, 1062–1074 (1983).
- Jørgensen, B. B. & Cohen, Y. *Limnol. Oceanogr.* **22**, 657–666 (1977).
- Des Marais, D. J. et al. *Nature* **359**, 605–609 (1992).
- Holland, H. D. in *Early Life on Earth* (ed. Bengtson, S.) 237–244 (Columbia University Press, New York, 1994).
- Canfield, D. E. et al. *Chem. Geol.* **54**, 149–155 (1986).
- Des Marais, D. J. et al. *Chem. Geol.* **65**, 341–358 (1987).
- Cline, J. D. *Limnol. Oceanogr.* **14**, 454–458 (1969).

ACKNOWLEDGEMENTS. This work was supported by the Max Planck Society, the Deutsche Sonderforschungsbereich, and NASA. We thank B. Thamdrup and B. B. Jørgensen for helpful discussion and comments during the preparation of the manuscript, and D. Des Marais for the PPRG stromatolite samples.

CORRESPONDENCE should be addressed to D.E.C. (e-mail: don@postgate.mpi-mm.uni-bremen.de).

Two-way exchange between the Easter mantle plume and the Easter microplate spreading axis

K. M. Haase^{*†‡}, C. W. Devey^{*} & S. L. Goldstein[†]

^{*} Geologisch-Paläontologisches Institut der Universität Kiel, Olshausenstrasse 40, D-24118 Kiel, Germany

[†] Max-Planck-Institut für Chemie, Postfach 3060, D-55020 Mainz, Germany

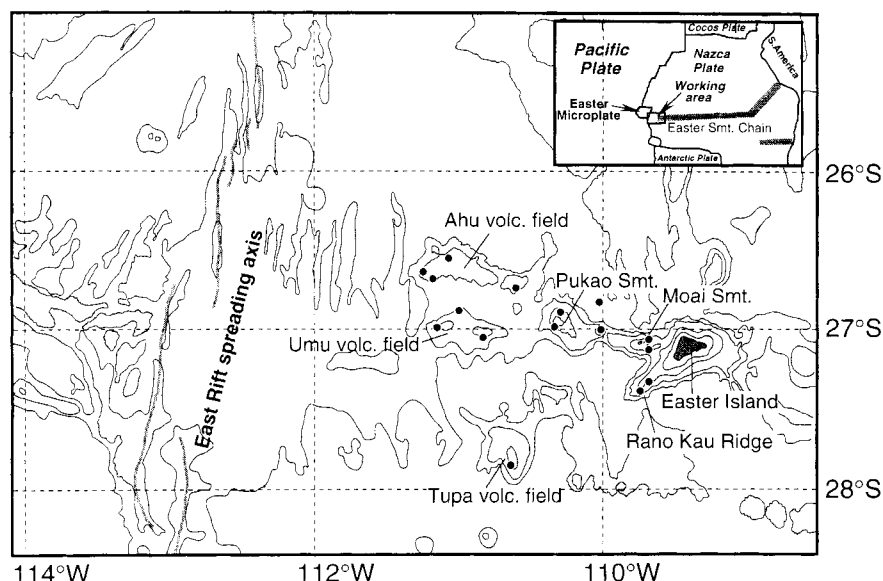
THE conventional model whereby plume volcanism forms linear age-progressive volcanic chains, with the youngest activity occurring nearest a spreading axis (at a 'hotspot'), has been challenged for the Easter seamount chain^{1–4}. Whereas early work suggested the existence of a linear melting anomaly (a 'hotline')^{1,2}, more recent studies^{3,4} have proposed a hotspot near Salas y Gomez island, connected with the Easter microplate spreading system by an ~800-km-long, volcanically active plume channel. Here we use geochemical, geological and geochronological data to argue that the hotspot lies close to Easter Island. Moreover, new isotopic data for lavas from the seamount chain provide evidence for bi-directional flow between the spreading axis and the plume, thus supporting geophysical and fluid-dynamical models of mantle flow in a plume/spreading axis system^{5–7}. Material balance and flux considerations show the Easter plume to be weak and cool compared with those beneath larger features such as Iceland, Hawaii and the Galápagos islands.

Easter Island lies at the western end of the Easter seamount chain (ESC), ~350 km east of the Easter microplate East Rift spreading axis (Fig. 1). Between Easter Island and the East Rift, on 1.5–4 Myr-old sea floor⁸ and occupying a region of 40,000 km², lie Pukao and Moai seamounts, and the Ahu, Umu and Tupa volcanic fields, and a 150-km-wide non-volcanic region (Fig. 1)^{9,10}. Pukao and Moai are older than the volcanic fields, although young glassy flows occur on their flanks^{9,10}. It has been suggested that the volcanic fields mark the 'Ahu Hotspot'^{9,11}. New ⁴⁰Ar–³⁹Ar data indicate an age progression along the ESC, with the youngest ages occurring on and west of Easter Island⁴. Despite these ages, some plate-tectonic reconstructions place the hotspot near Salas y

Gomez island, which is 400 km east of Easter^{4,12}, whereas others place it west of Easter¹³. Geoid studies have favoured a location near Easter^{14,15}. Schilling and co-workers^{3,16,17} favour a location near Salas y Gomez, because average Pb and Sr isotopic ratios are higher than for Easter, and it fits better with global correlations between distances to hotspots, lengths of plume-related geochemical anomalies on ridges, and sizes of ridge bathymetric anomalies.

We suggest that current data favour a plume location near Easter Island, based on several observations. (1) A regression line through ESC age data (Fig. 2a) crosses 0 Myr close to Easter Island. Of the 28 analysed Easter lavas, 25 show ages <0.9 Myr. These include the stratigraphically oldest samples, whose ages of 0.5–0.7 Myr (ref. 18) are much younger than the ~1.7 Myr activity of the more deeply eroded Salas y Gomez^{2,19}. The plate-tectonic reconstruction suggesting a hotspot location near Salas y Gomez⁴ assumes that the absolute velocity of the Nazca plate slowed to 4.7 cm yr⁻¹ ~5 Myr ago, and that dated ESC lavas represent the main volcanic stages. Although the typical lifetime of volcanoes (~3 Myr) adds a large uncertainty to age–distance relationships^{4,11}, available age data for the Easter, Galapagos and Juan Fernandez chains suggest the Nazca plate velocity was higher during the last 5 Myr. (2) The volcanoes between Easter and the East Rift are larger (>1,500 m high)^{9,10} than normal off-axis East Pacific Rise volcanoes (generally <800 m; ref. 20). The large magma volume indicates a strong thermal anomaly. High mantle temperatures in the Easter region are further indicated by lava compositions requiring higher degrees of melting than at Salas y Gomez³. Moreover, young volcanic edifices near Salas y Gomez are much smaller than those near Easter²¹. (3) Isotope ratios of Easter region samples showing the strongest plume signature (low ¹⁴³Nd/¹⁴⁴Nd, high ⁸⁷Sr/⁸⁶Sr and Pb isotopes) are similar to others from the ESC, including Salas y Gomez (Figs 2, 3), indicating that isotopic compositions of the plume source over the past 30 Myr has been nearly constant. The higher average Pb isotope ratio for Salas y Gomez has been cited as a strong argument for it as the location of the plume³, but this conclusion is based on three samples from Salas y Gomez and many from Easter and therefore has no statistical significance. Moreover, the lower Pb isotope ratios in some Easter, Moai and Pukao lavas show the influence of Easter microplate mid-ocean-ridge basalt (MORB) in their mantle source (Fig. 3), reflecting the location much closer to the spreading axis (Fig. 1). Thus, it is not the average isotopic composition which is relevant to the question of whether the plume is located beneath Easter Island. Rather, the important observation is that the pure plume signature is present

FIG. 1 Map of the working area, outlining the volcanic edifices of the Easter hotspot and the East Rift spreading axis redrawn after ref. 8. The filled circles show the sample stations of this study, and the major young volcanic edifices are labelled.



in many Easter lavas. (4) The empirical relationships between spreading-axis geochemical anomaly width, bathymetry, and the axis-to-hotspot distance are too imprecise to locate the hotspot exactly. Furthermore, the East Rift geochemical anomaly appears to be ~450 km long (24° 30' S to 28° 30' S; Fig. 5) rather than 350 km (ref. 3), suggesting that the hotspot lies closer to the Rift than Salas y Gomez.

Our new isotopic data (see supplementary information for tabulated data) support the suggestion³ of large-scale binary mixing between the Easter hotspot and the Easter microplate spreading axes. Moreover, isotopically 'enriched' lavas from the East Rift nearest Easter Island have values identical to those of volcanic-fields lavas (Fig. 3), implying shallow flow of MORB material into the intraplate region and thus bidirectional transfer

of material between spreading axis and plume. Lavas showing undiluted plume signatures erupt close to others from mixed sources, indicating that the mixing occurs at shallow levels rather than during plume ascent. The volcanic fields lavas and tholeiites from the spreading segment with the strongest plume influence show remarkable isotopic homogeneity, indicating efficient mixing at low viscosity.

Volcanic-fields tholeiites generally show a stronger heavy rare-earth fractionation than East Rift MORB, although they are isotopically similar. At appropriate degrees of melting, the only common residual mantle mineral that significantly fractionates Dy/Yb is garnet. Dy/Yb ratios increase with distance from the spreading axis (Fig. 4), implying increasing influence of garnet during melting. This reflects increasing lithospheric thickness as the plate moves away from the East Rift, which cuts the melting column from above, combined with increasing mantle temperatures approaching the plume, causing melting to begin at greater depths²².

Our data allow the first geochemical mass balance of the bidirectional material flow between a plume and a spreading axis. The rate of intraplate erupted magmas from the volcanic fields, Easter Island and the seamounts is $0.16 \text{ m}^3 \text{ s}^{-1}$ ($\sim 15 \times 10^{12} \text{ m}^3$ erupted over $\sim 3 \text{ Myr}$). A ridge extrusive/intrusive magma ratio of 1/6 (ref. 23) yields a magma production rate of $0.96 \text{ m}^3 \text{ s}^{-1}$, compared to $5.1 \text{ m}^3 \text{ s}^{-1}$ and $7.6 \text{ m}^3 \text{ s}^{-1}$ for Hawaii and Iceland²⁴, respectively. The Galápagos hotspot shows a higher production rate of $\sim 1.86 \text{ m}^3 \text{ s}^{-1}$ of magma (assuming 10 km of plume-generated crust²⁵, a diameter of 150 km and an age of 3 Myr; ref. 26). At 10% melting²², the upwelling mantle volume of the Easter plume would be $9.6 \text{ m}^3 \text{ s}^{-1}$ if the intraplate magmas were of pure plume origin. The intraplate lavas average $\sim 60\%$ plume material (Fig. 3), implying a plume input of $5.8 \text{ m}^3 \text{ s}^{-1}$. The spreading axes between 24° 30' S and 28° 30' S produce $\sim 11.4 \text{ m}^3 \text{ s}^{-1}$ of melt (8 km crustal thickness, 100 km Myr^{-1} spreading rate), implying a mantle volume of $87.6 \text{ m}^3 \text{ s}^{-1}$ at 13% melting²². However, depleted MORB material is contaminated to varying degrees by the plume (Fig. 3). We estimate a $\sim 30\%$ plume contribution to axis magmas ($26.3 \text{ m}^3 \text{ s}^{-1}$) (Fig. 3). These considerations indicate that the Easter plume volume which melts

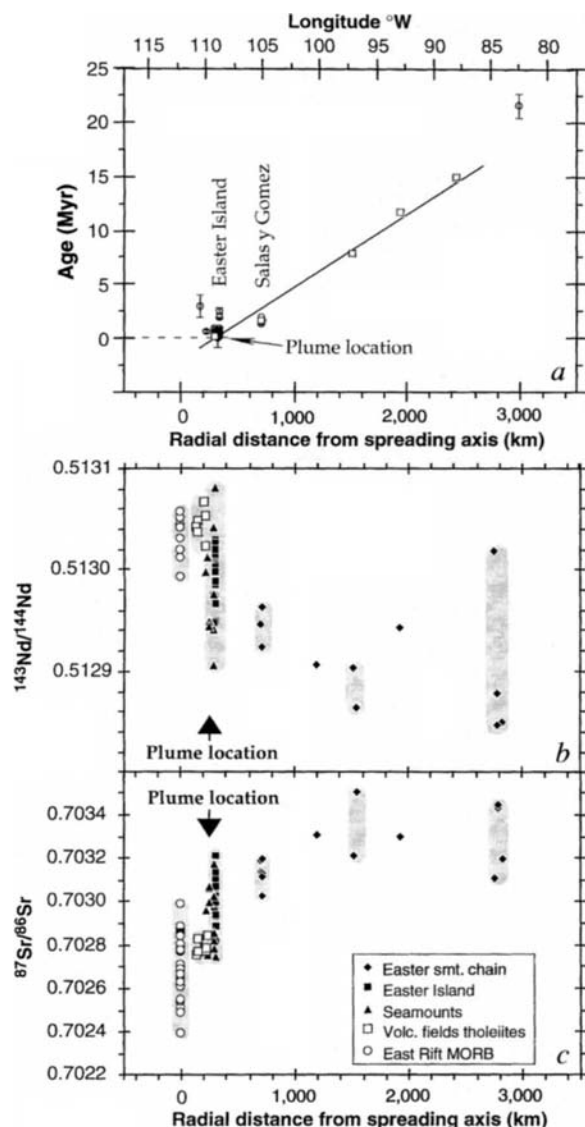


FIG. 2 a, Compilation of age data for the Easter seamount chain (ESC) and a calculated regression line, b, Nd isotopes and c, Sr isotopes, for lavas erupted along the ESC. The young ages of the Easter Island volcanism and the low-Nd but high-Sr isotopes are consistent with the plume being located there. The regression was calculated after the method of York³³ using only the data represented by the squares to ensure balanced statistics. The error of this regression calculation is $\pm 50 \text{ km}$ (2σ), whereas a regression of only the three samples between 5 and 15 Myr suggests the hotspot is 120 km east of Easter Island and 280 km west of Salas y Gomez but has a large 2σ error of 1,035 km. Data from this Letter and refs. 4, 34, 34, 35.

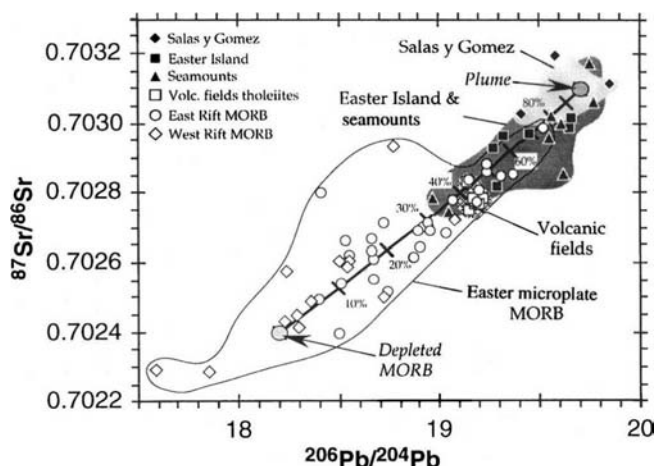


FIG. 3 $^{87}\text{Sr}/^{86}\text{Sr}$ versus $^{206}\text{Pb}/^{204}\text{Pb}$ with a mixing curve suggesting input of variable amounts of enriched plume material into MORB mantle. The mixing line is calculated assuming a MORB mantle containing 0.045 p.p.m. Pb and 13 p.p.m. Sr with plume mantle containing 0.1 p.p.m. Pb and 20 p.p.m. Sr calculated from 10% and 5% batch melting of respective basalts. Tick marks show percentage of plume material in the mix. As isotopic compositions of the endmembers we chose that of depleted Easter microplate MORB and of extreme Easter hotspot lavas, respectively. Note the similarity of compositions of many Easter Island and seamount samples to Salas y Gomez lavas. Data from this paper and refs 3, 17.

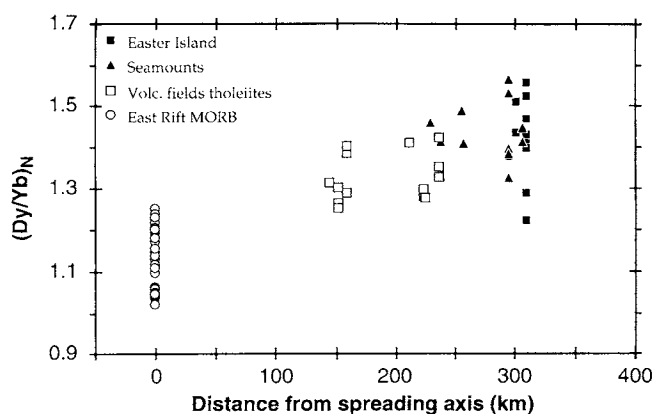


FIG. 4 Chondrite-normalized Dy/Yb ratio versus the distance from spreading axis. The increasing fractionation of the two heavy rare-earth elements suggests the increasing importance of residual garnet during melting as the lithosphere becomes older and thicker.

is $32.1 \text{ m}^3 \text{ s}^{-1}$, lower than some geophysical estimates of the total Easter plume flow (for example, $140 \text{ m}^3 \text{ s}^{-1}$; ref. 27) but comparable to others (for example, $41.5 \text{ m}^3 \text{ s}^{-1}$; ref. 28).

Our geochemical results support recent geophysical theoretical models and laboratory experiments for plume/spreading axis interactions. Passive flow of MORB with the lithosphere and buoyant flow of plume into the axis as described here (Fig. 5) was predicted by fluid-dynamical models⁵⁻⁷ and has recently been

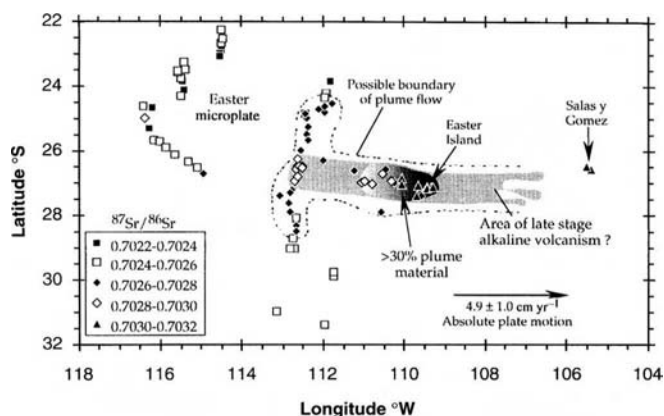


FIG. 5 Geographical plot of $^{87}\text{Sr}/^{86}\text{Sr}$ data and a model for the Easter hotspot/Easter microplate volcanism. The shading outlines the occurrence of lavas with large input of radiogenic material (Fig. 2) whereas the stippled line indicates the possible shape of the plume flow. The main plume channel feeds the spreading segment closest to Easter Island and symmetrical along-axis flow feeds the northern and southern segments. Voluminous young volcanism containing more than 30% plume material erupts above the plume centre, whereas small volumes of low-degree melts form further east above the sheared plume. Data from this work and refs 17 and 36, and plate velocity from ref. 37.

modelled rigorously by Sleep²⁹. At Easter, relatively low magma production rates, a low degree of partial melting for most of the intraplate magmas (perhaps 10%; ref. 22), and the small bathymetric anomaly ($\sim 0.4 \text{ km}$ above normal depth³⁰) compared to other plumes indicate a low excess temperature of $\sim 100^\circ \text{C}$. This appears to be further reflected in a weak seismic tomographic signature³¹ and in the fluid-dynamical behaviour⁷. Experiments indicate that weak plumes preferentially form channels toward the spreading axis and that plumes are deflected in the direction of plate movement⁷. Dragging of plume material by the lithosphere may be a cause of young alkalic volcanism as at Salas y Gomez^{21,32}. The geochemical influence of the plume on the East Rift between 24°S and 28°S (refs 3, 16) indicates along-axis flow of plume material (Fig. 5). The 150-km-wide volcanic gap between the volcanic fields and the spreading axis indicates fluid-dynamical focusing of magma into the spreading centre within this region. Whereas previous studies have concentrated on the plume flux into the spreading axis, the present work shows that a counterflow of MORB material occurs and that mixing of plume and MORB-sources is very efficient. □

Received 25 October 1995; accepted 30 May 1996.

- Bonatti, E. & Harrison, C. G. A. *Nature* **263**, 402–404 (1976).
- Bonatti, E. et al. *J. geophys. Res.* **82**, 2457–2478 (1977).
- Hanan, B. B. & Schilling, J.-G. *J. geophys. Res.* **94**, 7432–7448 (1989).
- O'Connor, J. M., Stoffers, P. & McWilliams, M. O. *Earth planet. Sci. Lett.* **136**, 197–212 (1995).
- Sleep, N. H. *Geophys. J. R. astr. Soc.* **91**, 1–11 (1987).
- Phipps Morgan, J. & Smith, W. H. F. *Nature* **359**, 524–527 (1992).
- Kincaid, C., Ito, G. & Gable, C. *Nature* **376**, 758–761 (1995).
- Naar, D. F. & Hey, R. N. *J. geophys. Res.* **96**, 7961–7993 (1991).
- Hagen, R. A., Baker, N. A., Naar, D. F. & Hey, R. N. *Mar. geophys. Res.* **12**, 297–315 (1990).
- Maia, M., Dehghani, G. A., Diamant, M., Francheteau, J. & Stoffers, P. *Geophys. Res. Lett.* **118**, 177–185 (1994).
- Naar, D. F. & Hey, R. N. in *Evolution of Mid Ocean Ridges* (ed. Sinton, J. M.) 9–30 (Geophysical Monogr. Am. Geophys. Union, Washington DC, 1989).
- Okal, E. A. & Cazenave, A. *Earth planet. Sci. Lett.* **72**, 99–116 (1985).
- Pilger, R. H. & Handschumacher, D. W. *Geol. Soc. Am. Bull.* **92**, 437–446 (1981).
- Fleitout, L. & Moriceau, C. *Geophys. J. Int.* **110**, 6–82 (1992).
- Maia, M., Dehghani, G. A., Diamant, M., Francheteau, J. & Stoffers, P. *Geophys. Res. Lett.* **21**, 2527–2530 (1994).
- Schilling, J.-G., Sigurdsson, H., Davis, A. N. & Hey, R. N. *Nature* **317**, 325–331 (1985).
- Fontignie, D. & Schilling, J.-G. *Chem. Geol.* **89**, 209–241 (1991).
- Kaneoka, I. & Katsui, Y. *Bull. Volcan. Soc. Japan* **30**, 33–36 (1985).
- Clark, J. G. & Dymond, J. J. *Volcan. geotherm. Res.* **2**, 29–48 (1997).
- Scheirer, D. S. & Macdonald, K. C. *J. geophys. Res.* **100**, 2239–2259 (1995).
- Naar, D. F. and the Scientific Party Eos **74**, 672 (1993).
- Haase, K. M. & Devey, C. W. *Earth planet. Sci. Lett.* **137**, 129–143 (1996).
- Crisp, J. A. *J. Volcan. geotherm. Res.* **20**, 177–211 (1984).
- White, R. S. *Phil. Trans. R. Soc. Lond.* **A342**, 137–153 (1993).
- Feighner, M. A. & Richards, M. A. *J. geophys. Res.* **99**, 6711–6729 (1994).
- White, W. M., McBirney, A. R. & Duncan, R. A. *J. geophys. Res.* **98**, 19533–19563 (1993).
- Sleep, N. H. *J. geophys. Res.* **95**, 6715–6736 (1990).
- Schilling, J.-G. *Nature* **352**, 397–403 (1991).
- Sleep, N. H. *J. geophys. Res.* (in the press).
- Schroeder, W. J. *Geophys. Res.* **89**, 9873–9883 (1984).
- Zhang, Y. S. & Tanimoto, T. *J. geophys. Res.* **98**, 9793–9823 (1993).
- Poreda, R. J., Schilling, J.-G., Batiza, R. & Naar, D. *Eos* **74**, 672 (1993).
- York, D. *Earth planet. Sci. Lett.* **5**, 320–325 (1969).
- Cheng, Q. C. thesis, Univ. California, San Diego (1989).
- MacDougall, J. D. & Lugmair, G. W. *Earth planet. Sci. Lett.* **77**, 273–284 (1986).
- Haase, K. M. thesis, Christian-Albrechts-Univ., Kiel (1995).
- Gripp, A. E. & Gordon, R. G. *Geophys. Res. Lett.* **17**, 1109–1112 (1990).

SUPPLEMENTARY INFORMATION. Available on Nature's World-Wide Web site <http://www.nature.com>. Paper copies are available from Mary Sheehan at the London editorial office of Nature.

ACKNOWLEDGEMENTS. We thank the captain and the crew FS Sonne and P. Stoffers for providing the samples and supporting the study. We also thank D. Garbe-Schönberg and S. Raczek for their help during the analytical work, and N. Sleep for reviews. K.M.H. was supported by the DFG through the Graduiertenkolleg 'Dynamik Globaler Stoffkreisläufe', this work was supported by a BMFT grant to P. Stoffers.

CORRESPONDENCE should be addressed to K.M.H. (e-mail: haase@geobar.mpgch-mainz.mpg.de).

Eel-like swimming in the earliest ichthyosaurs

R. Motani*†, You H.‡ & C. McGowan*†

* Department of Zoology, University of Toronto, Toronto, Ontario M5S 1A1, Canada

† Department of Paleobiology, Royal Ontario Museum, Toronto, Ontario M5S 2C6, Canada

‡ Institute of Vertebrate Paleontology and Paleoanthropology, PO Box 643, Beijing 100044, China

ICHTHYOSAURS are extinct marine reptiles, probably belonging to the Diapsida¹, that ranged from the Early Triassic to Late Cretaceous^{2,3}. Post-Triassic ichthyosaurs achieved the highest level of aquatic adaptation among reptiles⁴, with a streamlined body, lunate tail and a dorsal fin, features exemplified today by thunniform (tuna-like) fishes. However, little is known of how such a body plan evolved from a terrestrial diapsid. Here we report the most complete specimen of the oldest known ichthyosaur, *Chensaurus*, representing a transition between the two body plans. The specimen, which has a partial skin impression, has a small caudal fin, a long and narrow body, and a high presacral vertebral count. These features all suggest an anguilliform swimming mode. Later ichthyosaurs retained the high vertebral count, but overcame the high swimming costs of this plesiomorphy, achieving a rigid thunniform bauplan by evolving discoidal vertebrae, and a deep fusiform body. *Chensaurus* therefore seems to be an evolutionary intermediate between the shorter-bodied terrestrial stock from which the group evolved, and advanced thunniform ichthyosaurs.

The specimen was collected in 1989 from the Lower Triassic (Spathian) of Anhui Province, China, about 50 km southwest of the type locality of *Chensaurus*^{5,6}, which is also Spathian⁷. Based on overall similarities, the specimen is tentatively identified as *Chensaurus chaioxianensis* (Chen) 1985 (ref. 5). Its most remarkable feature is its slender trunk region (Fig. 1a), rare among ichthyosaurs. This slenderness is not a post-mortem artefact, because the gastralia are preserved *in situ*, the articulated series lying parallel to the vertebral column. Also noteworthy is that the body outline is partially preserved in the dorsal region. The outline of the caudal fin is the best preserved, located immediately dorsal to a change in orientation of the neural spines (Fig. 1a).

Sharks and ichthyosaurs are similar in that their vertebral columns continue into one of the caudal fin lobes: the upper and lower lobes, respectively⁸. Also, they both have high precaudal vertebral counts (usually 60–110), in contrast to scombrid fishes (about 40), and cetaceans (usually 40–60). Sharks evolved several body forms, some of which are also found in ichthyosaurs (Fig. 1). Because of these similarities, sharks provide the best analogue for ichthyosaurs in overall body shape and locomotion, although differing in details.

Vertebrates that swim by lateral undulations of the body may be described as anguilliform, sub-carangiform, carangiform, and thunniform, according to the proportion of the body utilized for the propulsion (highest in anguilliform and lowest in thunniform)⁹. These modes are also associated with body shape, ranging from the elongate and flexible anguilliform swimmers to the deep and more rigid-bodied thunniform ones, as noted among teleosts and sharks¹⁰. Sharks range from being anguilliform to thunniform⁸, though some authorities avoid these categories, preferring to use informal groups defined on shape¹¹. The variation in body and tail shape among sharks is considerable, and we used two indices to quantify this: fineness ratio (precaudal length/body height) and tail H/L (tail height/length; Fig. 2a). These indices are positively correlated ($r = 0.70$; Fig. 2a) and show a trend from the more anguilliform sharks, as exemplified by scyliorhinids, to the more thunniform ones, like lamnids (Fig. 2b). Two ichthyosaur genera were added to the data: *Stenopterygius* (Fig. 1d), a typical

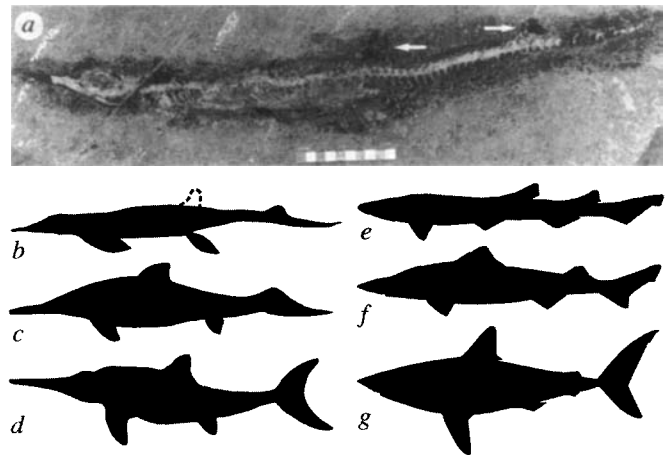


FIG. 1 Variation in body form among ichthyosaurs and sharks. a, The new specimen of *Chensaurus chaioxianensis*, Wuwei Cultural Relic Administrative Institute, Anhui Province, China (WCRAI 313). The fossil has been cleaved along the sagittal plane, and is disposed on two slabs. The left-hand arrow marks the position of a feature of uncertain identity (dorsal fin or hind fin), the right-hand arrow marks the anterior end of the caudal fin. The body outline is the narrow black zone lying close to the dorsal edge of the vertebral column. It is easiest to see in the region between the arrows. The dark zone lying ventral to the vertebral column is a recent artifact (a separator, used during mould making). A fault in the cervical region displaced the skull and the first several vertebrae relative to the body. Scale bar, 10 cm. b, reconstruction of WCRAI 313. Paired fins added from a smaller specimen, scaled to the appropriate size. The questionable feature is depicted by the broken line. Note the narrow body and small caudal fin. c, Reconstruction of *Mixosaurus cornalianus*, modified from ref. 23. The trunk is fusiform but the caudal fin is similar to that of *Chensaurus*. d, *Stenopterygius quadricissus*, Paleontologiska Museet, Uppsala Universitet (PMU R158). Body outline preserved as a carbonaceous film, showing deep fusiform body, lunate tail and dorsal fin. e, *Asymbolus vincenti*, a scyliorhinid shark with a body outline resembling *Chensaurus*. f, *Centrophorus harrisoni*, a squalid shark similar in shape to *Mixosaurus*. g, *Lamna nasus*, a lamnid shark whose body plan is similar to *Stenopterygius*. e–g modified from ref. 24.

post-Triassic form, interpreted as being adapted for fast cruising^{8,12,13}, and *Chensaurus* (Fig. 1b). *Stenopterygius* groups with the lamnid sharks, whereas *Chensaurus* lies at the extreme end of the scyliorhinid distribution, suggesting that it was anguilliform, and a forerunner of the advanced thunniform ichthyosaurs.

Anguilliform swimming requires body flexibility, which is enhanced by high vertebral counts. The presacral count of *C. chaioxianensis* is about 40, some 50 per cent higher than that of most limbed terrestrial amniotes, both living and extinct^{14,15}. Most later ichthyosaurs have 40–50 presacrals^{16–18}, except for long-bodied shastasaurids with approximately 65 (refs 19,20). The presacral count of *C. chaioxianensis* is therefore already within the range of typical ichthyosaurs. This suggests that a high presacral count appeared early in ichthyosaur evolution, as an adaptation for anguilliform swimming, and was retained in later forms.

The optimum efficiency of thunniform swimmers is achieved by a stiff body, limiting lateral propulsive movements to the caudal fin. Body stiffness is enhanced by restricting the degree of flexion between adjacent vertebrae. For ichthyosaurs, which have amphicoelous vertebrae, intervertebral flexion was probably largely a function of the thickness and compliance of the intervertebral discs. From simple geometry, the maximum angular displacement between adjacent vertebral centra decreases with increasing diameter, other dimensions remaining the same. *C. chaioxianensis*, and other Early Triassic ichthyosaurs, have cylindrical centra, but they became taller and wider in later forms, culminating in the

FIG. 2 The correlation between body shape and tail shape in sharks. *a*, 95% confidence ellipse fitted to data for 94 species, belonging to 14 families, ($r = 0.70$, $**P < 0.01$; r for the population estimated at $0.58\text{--}0.80$, $*P < 0.05$). Data were obtained by taking measurements from published figures^{24,25}. The boundary between precaudal and caudal regions was determined according to a published method²². *b*, As *a*, but for ease of comparison only three shark families are depicted. It was not plausible to fit a 95% confidence ellipse to the lamnid data because of the small sample size. Note that the ichthyosaurs, *Chensaurus* and *Stenopterygius*, lie at the two extremes of the shark distributions.

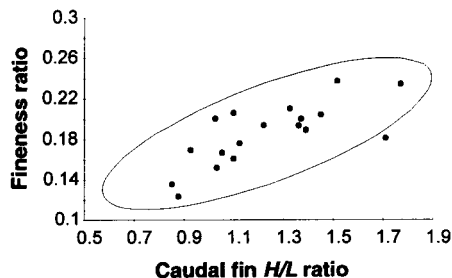
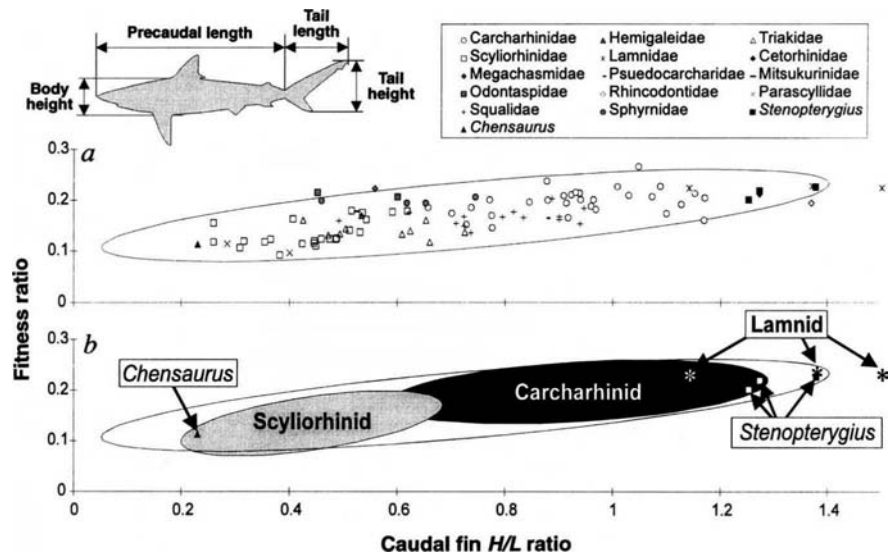


FIG. 3 The correlation between body shape and centrum shape in sharks. Data for 18 species belonging to 6 families were obtained from the literature^{22,24}, ($r = 0.74$, $**P < 0.01$; r for the population estimated at $0.41\text{--}0.91$, $*P < 0.05$). Two species, *Prionace glauca* (blue shark) and *Eusphyra blochii* (winghead shark), are unusual in precaudal count, hence were not included: blue shark has about 146 precaudals, almost 52 per cent more than average carcharhinid sharks, and winghead shark has about 50, nearly 47 per cent less than average hammerhead sharks. Measurements for ichthyosaurs (referred to in text) were taken for posterior dorsal vertebrae; those for sharks were for penultimate monospondylous vertebrae²².

discoidal centra typical of most ichthyosaurs. The depth of the centrum can be expressed by the ratio of centrum height to length (H/L). The ratio has values of 0.9 for *C. chaoyangensis*, 2.0 for *M. cornalianus* (Middle Triassic) and 2.5 for *S. quadriscissus* (Lower Jurassic). Ichthyosaurs therefore seem to have overcome the problem of retaining their high presacral counts—the antithesis of rigidity—by evolving deep discoidal centra. Associated with this change was a deepening of the body. Riess²¹ described *Mixosaurus* as an anguilliform swimmer, but his argument only establish that it was an axial swimmer, which accords with our hypothesis.

The only living animals with comparable centra are sharks, whose fossilized vertebrae are sometimes confused with those of

ichthyosaurs. Sharks, as noted, usually have precaudal counts similar to ichthyosaurs. Significantly, the deeper bodied sharks tend to have the deepest centra (Fig. 3), with the exception of those with unusual precaudal counts. Thus the centrum H/L ratio for scylliorhinid sharks average about 0.8, compared with 2.0 for lamnids²². This supports our contention that the evolution of a deep, fusiform body, typical of post-Triassic ichthyosaurs, was predicated upon the evolution of discoidal vertebrae. Anthracosaurs also have discoidal vertebrae, but their body plan is not comparable to that of ichthyosaurs: presence of both intercentra and pleurocentra resulted in about 80 joints in the presacral region, contributing flexibility to their presumed anguilliform locomotion. □

Received 25 March; accepted 31 May 1996.

- Massare, J. A. & Callaway, J. M. *Geol. Soc. Am. Bull.* **102**, 409–416 (1990).
- Baird, D. *Mosasur* **2**, 129–133 (1984).
- Callaway, J. M. & Massare, J. A. *Neues Jb. Geol. Palaont. Mh.* **178**, 37–58 (1989).
- Carroll, R. L. *Spec. Pap. Palaont.* **33**, 145–155 (1985).
- Chen, L.-Z. *Reg. Geol. China* **15**, 139–146 (1985).
- Mazin, J.-M., Suteethorn, V., Buffetaut, E., Jaeger, J.-J. & Helmcke-Ingavat, R. *C. r. hebdomadaire. Séanc. Acad. Sci., Paris* **313**, 1207–1212 (1991).
- Wang, G. X. *Marine Facies of Triassic System of Anhui 1–73* (Anhui Science and Technology, Hefei, 1984).
- McGowan, C. *Palaentology* **35**, 555–570 (1992).
- Webb, P. W. & Blake, R. W. in *Functional Vertebrate Morphology* (eds Hildebrand, M., Bramble, D. M., Liem, K. F. & Wake, D. B.) 110–128 (Harvard Univ. Press, Cambridge, MA, 1985).
- Webb, P. W. & Keyes, R. S. *Fish. Bull. F.A.O.* **89**, 803–812 (1982).
- Thomson, K. S. & Simanek, D. E. *Am. Zool.* **17**, 343–354 (1977).
- Webb, P. W. *Am. Zool.* **28**, 709–725 (1988).
- Massare, J. A. *Paleobiology* **14**, 187–205 (1988).
- Romer, A. S. *The Osteology of the Reptiles 1–772* (Univ. Chicago Press, 1956).
- Hoffstetter, R. & Gasc, J.-P. in *Biology of the Reptilia Vol. 1* (eds Gans, C., Bellairs, A. d'A. &

- Parsons, T. S.) 201–310 (Academic, New York, 1969).
- McGowan, C. *Life Sci. Contr. R. Ont. Mus.* **97**, 1–37 (1974).
- McGowan, C. *Life Sci. Contr. R. Ont. Mus.* **100**, 1–30 (1974).
- McGowan, C. *Palaentogr. am.* **A166**, 93–135 (1979).
- Merriam, J. C. *Mem. Univ. Calif.* **1**, 1–196 (1908).
- Kosch, B. F. J. *Vertebr. Paleontol.* **10**, 512–514 (1990).
- Riess, J. *Palaentogr. am.* **A192**, 93–155 (1986).
- Springer, V. G. & Garrick, J. A. F. *Proc. U.S. natn. Mus.* **116**, 73–96 (1964).
- Kuhn-Schwyder, E. *Neujahrsblatt, Die Triasfauna der Tessiner Kalkalpen 1–119* (Naturforschenden Gesellschaft, Zurich, 1974).
- Last, P. R. & Stevens, J. D. *Sharks and Rays of Australia 1–513* (CSIRO Australia, 1994).
- Hauff, B. & Hauff, R. B. *Das Holzmadenbuch 1–136* (Repro-Druck, Fellbach, 1981).

ACKNOWLEDGEMENTS. We thank Z. Dong and Y. Tomida for arranging the study. X. Ye, F. He and A. He for help in Wuwei. A. Baker, H.-D. Sues, M. A. Taylor, and P. W. Webb for useful comments. This project was supported by Fujiwara Natural History Foundation (Tokyo) grant to R.M. and National Science and Engineering Research Council grant to C.McG.

CORRESPONDENCE and requests for materials should be addressed to R.M. (e-mail: rmtani@utcc.utoronto.ca).

A new specimen of *Ankarapithecus meteai* from the Sinap Formation of central Anatolia

Berna Alpagut*, Peter Andrews†, Mikael Fortelius‡,
John Kappelman§, İlhan Temizsoy||, Hürkan Çelebi*
& William Lindsay†

* Paleontoloji, Ankara Üniversitesi, Ankara, Turkey

† Natural History Museum, Cromwell Road, London SW7 5BD, UK

‡ Finnish Museum of Natural History, PO Box 11,
FIN-00014 University of Helsinki, Finland

§ Department of Anthropology, The University of Texas at Austin,
Texas 78712-1086, USA

|| Anadolu Medeniyetleri Müzesi, Ankara, Turkey

HOMINOID fossils from the Middle and Late Miocene are exceedingly rare, yet such material is necessary for determining hominoid phylogeny. We report here the discovery of a fossil hominoid partial skull from the Upper Miocene Sinap Formation^{1,2} of central Turkey that is the most complete known from the period of 18 to 3 Myr. Our fieldwork places the hominoid locality within a precisely dated geochronological and biostratigraphical framework³ that permits detailed comparisons with other fossil hominoids. Earlier discoveries of more fragmentary remains of *Ankarapithecus meteai* suggested affinities with the Asian hominoids *Sivapithecus* and *Pongo*⁴. This new and nearly complete specimen reveals a combination of facial, mandibular, and dental features including a relatively narrow interorbital region, extensive frontal and maxillary sinuses, moderately developed supraorbital tori, square orbits, robust mandibular corpus, and incisor heteromorphy that is not matched in any extant or fossil hominoid. This configuration of features seems to support its placement as a stem member of the great ape and human clade.

The hominoid partial skull and several postcranial elements (to be described elsewhere) were discovered at locality 12 (40° 15' 02" N, 32° 38' 58" E) on the southernmost hilltop of Delikayınca Tepe, a prominent ridge located about 2.5 km north of the village of Yassıören. Fossils have been recovered from the Sinap Formation for more than 40 years^{5,6}, and some of these can be placed within our precise stratigraphical framework that includes nearly 100 fossil localities in the Formation¹⁻³. Earlier discoveries include two specimens of fossil hominoid, the first being the type mandible of *Ankarapithecus meteai* (Ozansoy 1957)⁵, which came from our locality 8B, and a lower face (specimen MTA 2125) from locality 12 attributed to the same taxon⁴.

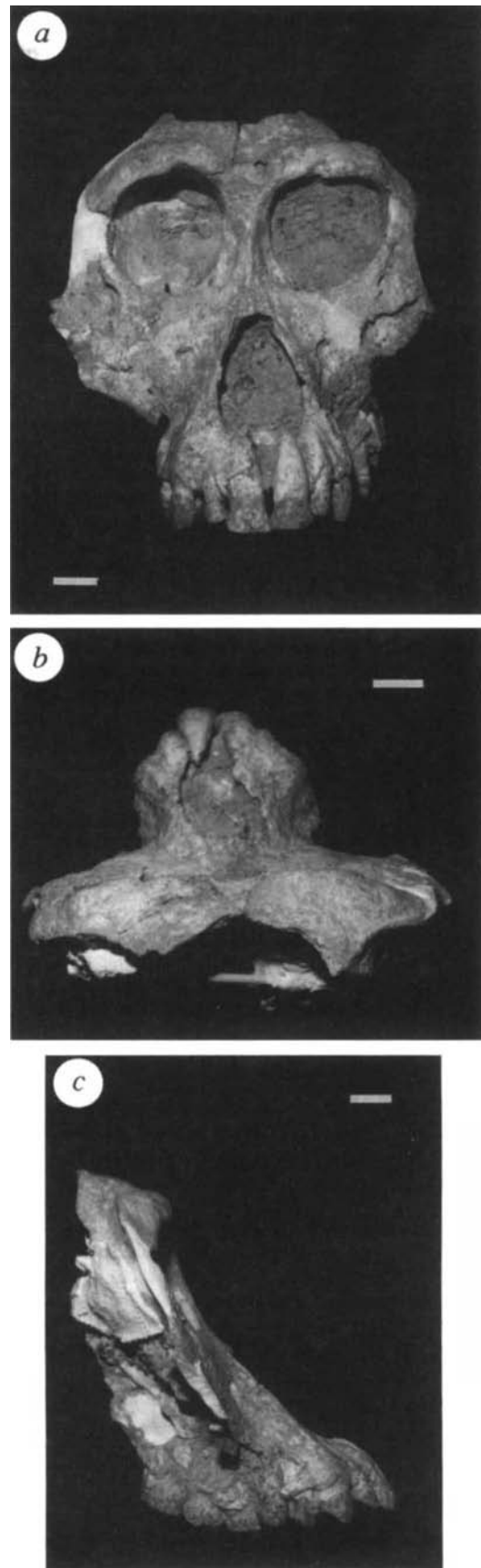


FIG. 1 The mandible and face were found about 20 cm apart at the same stratigraphic level. The fossils from locality 12 are generally well preserved although the surface bone of many specimens including the hominoids is sometimes heavily marked by tiny root traces, and postcranial elements of both large- and small-bodied taxa are sometimes preserved as associated and articulated specimens. There was probably little post-depositional transport, and it seems likely that initial burial in the encasing pebbly siltstone was rapid but shallow and within the root zone of the overlying and poorly developed paleosol. The face of AS95-500 includes most of the premaxilla, maxilla, zygomatics, lacrimals, and nasals, and portions of the sphenoids, palatines, and frontal including a portion of the right and left temporal lines. a, Frontal view; b, Superior view; c, Right lateral view. A small portion of the frontal process of the right zygomatic is missing just inferior to the zygomaticofrontal suture. This anatomy is preserved on the left side but minor breakage and distortion along the upper lateral margin of the frontal gives the left-eye orbit a more sharply angular appearance. In side view the facial profile slopes smoothly upward from the incisors to a more vertical inclination about midway between rhinion and nasion that continues to glabella. See Table 1 for measurements. The upper dentition is complete. The arcade of the upper

incisors through to the premolars is also perfectly preserved but all of the upper molars on both the right and left sides are displaced somewhat superiorly, offset laterally by about 1 cm to the left from the anterior tooth row. Measurements of the teeth are provided in Table 2. AS 95, Ankara Sinap 1995 field season. Specimens are housed in the Anatolian Museum of Civilization, Ankara. Scale bars, 1 cm.

These two localities are about 1 km apart but within a few metres of each other stratigraphically, and in the younger portion of the long normal of chron C5 of the geomagnetic polarity timescale with an interpolated age of about 9.8 Myr (refs 3, 7). This age is similar to those estimated for *Ouranopithecus macedoniensis* from Greece⁸ and *Dryopithecus laietanus* from Can Llobateres⁹, and near the middle of the age range for *Sivapithecus* from the Indian subcontinent^{10,11}.

The hominoid partial skull (specimen number AS95-500; Figs 1, 2) is an adult individual (fully erupted M₃ and M³ teeth) inferred to be female and conspecific with the earlier described specimens inferred to be male individuals. Revisions of hominoid systematics had treated *Ankarapithecus* as synonymous with the Asian genus *Sivapithecus*^{4,12,13}, but analysis of the new material from locality 12 demonstrates that it is distinct from that genus, as suggested recently after a reanalysis of the earlier material¹⁴.

The relative completeness of AS95-500 (Tables 1, 2) permits an assessment of many characteristics that are believed to be vital in determining the phylogeny of Miocene hominoids. Upper incisor heteromorphy is extreme in AS95-500 (Fig. 1a), as it is in *Pongo*, *Ouranopithecus* and *Sivapithecus*, but this characteristic may be of limited utility in separating taxa¹⁵. The orbits are about as broad as they are tall (Fig. 1a) as seen in the African hominoids and *Ouranopithecus*, and differ from the generally narrow and tall orbits of *Pongo* and *Sivapithecus*¹⁶. The interorbital region is as relatively narrow as that of *Pongo*¹⁶ but within the range of variation of *Pan paniscus*¹⁷. The configuration of sharply keeled nasals and anteriorly exposed lacrimals is somewhat similar to that of *Gorilla*, *Ouranopithecus*, and the much earlier *Afropithecus*. Computed tomography reveals an invasive frontal sinus that extends 24 mm laterally from the midline. The presence of such an extensive sinus along with a relatively narrow interorbital region contrasts with the morphology of *Pongo* and *Sivapithecus*¹⁶, but *Pan paniscus* does combine these features. It is not clear whether the frontal sinus arises from the ethmoidal air cells, as is the case of the African hominoids^{17,18}, or is an extension of the maxillary sinus as in *Pongo* and *Sivapithecus*¹⁹, but such an extensive frontal sinus is unknown for *Pongo*. The maxillary sinus is very invasive and extends into the root of the zygoma, which is flat and anteriorly flexed as in *Dryopithecus*, *Sivapithecus*, and *Pongo*²⁰. There is a pronounced canine fossa.

The naso-alveolar region is preserved in MTA 2125, but unfortunately the alveolar margin of the premaxillary region and anterior palate are heavily eroded in AS95-500 (Fig. 1b). It is evident, however, that the premaxilla was elongated and prognathous as in MTA 2125. Recent cleaning of MTA 2125 reveals a mildly stepped morphology with a large incisive canal and foramen¹⁴ in contrast to earlier descriptions^{21,22}. The frontal region is marked by moderately developed supraorbital tori (Fig. 1a) defined by a somewhat depressed glabellar region, and a poorly developed sulcus. This configuration differs from the generally inflated glabella of *Pan* and *Gorilla* that joins the two tori into one continuous torus. The angle of the frontal squama is most similar to that of *Ouranopithecus*²³ and *Dryopithecus*^{16,24}. The overall profile (Fig. 1c) is closer to that of the African hominoids than the concave facial profile and more vertically inclined frontal of *Pongo* and *Sivapithecus*.

Molar enamel seems to be thick and is probably within the

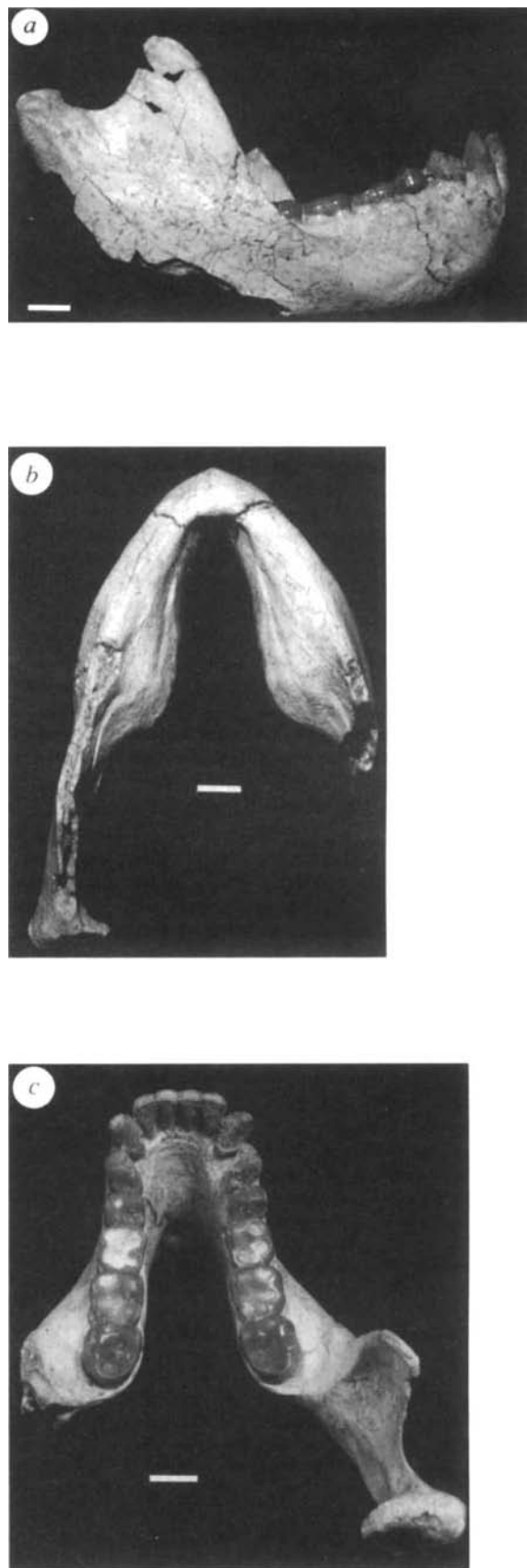


FIG. 2 The mandible is almost complete. a, Right side, slightly angled lateral view also showing the anterior portion of the left ramus. b, Inferior view. c, Occlusal view. The symphysis preserves the superior and inferior transverse tori. The corpus is robust and vertically fractured but not broken near the mental foramen. It is complete posteriorly beyond the oblique line on both the right and left sides with the internal surface showing an easily observable mylohyoid line. The ascending ramus is nearly complete on the right side and includes the complete anterior vertical margin, coronoid process, mandibular notch, and condyle. Only the inferior and posterior margins of the gonial angle of the right ramus are missing. The lower dentition is complete. The

mandibular arcade is undistorted and shows tooth rows that slightly diverge posteriorly. The size and the identical state of wear on the upper and lower dentitions including somewhat more advanced wear on the left side demonstrates that the face and mandible are from the same individual. Measurements are given in the Tables. Scale bars, 1 cm.

TABLE 1 Facial, maxillary and mandibular measurements of AS95-500

Nasal height	Est. 61.0
Alveolare to nasospinale	Est. 15.5
Nasal aperture height	Est. 25.1
Nasal aperture breadth	Est. 20.5
Nasal bone length	Est. 35.2
Orbital height	30.8
Orbital breadth	31.0
Zygomatic depth	29.3
Orbito-alveolar height	48.9
Minimum interorbital breadth	11.9
External biorbital breadth	Est. 87.2
Internal biorbital breadth	Est. 72.8
Lacrymal crests breadth	14.8
Lacrymal fossa breadth (right and left)	6.0
Lacrymal fossa length (right and left)	15.5
Post-orbital constriction	Est. 61.8
Palatal length	Est. 63.0
Palatal breadth at maxillary C	25.0
Palatal depth at maxillary C	Est. 5.5
Palatal depth at M ²	Est. 12.6
Alveolar breadth at maxillary C	42.9
Symphysis	Depth 36.2
	Thickness 16.0
Corpus at P ₄	Depth 27.5
	Thickness 14.4
Corpus at M ₃	Depth 26.2
	Thickness 26.1
Ramus	Notch height Est. 52.0
	Condylar height Est. 54.2
	Maximum breadth 44.8
	Length Est. 58.5
Bicondylar breadth	Est. 105.4
Mandibular angle	Est. 109°–121°

Lengths are given in mm. Est., estimated measurement.

TABLE 2 Dental measurements of AS95-500

Maxillary dentition		Left	Right
I ¹	MD	10.1	10.5
	LaLi	8.4	8.5
I ²	MD	6.3	6.4
	LaLi	6.3	6.3
C	Max.	9.0	9.2
	Tr	8.8	8.2
P ³	MD	7.2	7.4
	BuLi	12.3	12.4
P ⁴	MD	7.6	7.4
	BuLi	12.0	11.6
M ¹	MD	10.9	10.6
	BuLi	12.6	12.2
M ²	MD	12.1	12.3
	BuLi	13.7	13.6
M ³	MD	11.5	11.3
	BuLi	13.2	13.1
Mandibular dentition			
I ₁	MD	5.3	5.2
	LaLi	6.3	6.2
I ₂	MD	5.8	6.0
	LaLi	7.8	7.7
C	Max.	10.1	10.1
	Tr	6.4	6.4
P ₃	Max.	12.5	12.3
	Tr	7.1	6.7
P ₄	MD	7.5	7.6
	BuLi	9.9	9.5
M ₁	MD	11.0	10.0
	BuLi	10.6	10.7
M ₂	MD	12.2	12.3
	BuLi	12.1	12.0
M ₃	MD	13.6	13.8
	BuLi	12.7	13.0
	Talonid	12.0	12.3

Lengths are given in mm. MD, mesiodistal; LaLi, labiolingual; max., maximum; Tr, transverse; BuLi, buccolingual.

range of variation of most contemporary fossil hominoids, excluding *Dryopithecus*²⁵. The corpus of the mandible is extremely robust (Fig. 2), particularly in the region of the third molar, and it is similar in proportions to *Sivapithecus*¹⁹ and contrasts with the generally more shallow and gracile mandibles of *Dryopithecus*²⁶ and *Ouranopithecus*²⁷.

The mix of characters in AS95-500 indicates the need to reassess character definition and polarity in hominoid evolution. As noted above, some of these features are shared with *Pongo*, others are shared with the *Pan*, *Gorilla*, and sometimes *Dryopithecus*, whereas still others are shared with *Ouranopithecus*. Many of these features, such as a narrow interorbital region, flat zygomatic region, and elongated premaxilla have been linked to the complex of airorhynch^{28,29} as seen in *Pongo*. Other features, such as a supraorbital torus and frontal sinus, have been linked to klinorhynch²⁹ as seen in the African apes. Although such distinctions may be of some use in separating living taxa, the combination of these features in a single specimen indicates that the relative position of the face and cranial vault is more complex than a division into two extreme types.

Several conclusions can be drawn from this new specimen. If the characters shared with the Asian apes are accepted as homologous^{15,20,23}, the similarities with *Dryopithecus*, *Ouranopithecus* and in some cases the African apes would be homoplasies. However, if the similarities with *Dryopithecus* and *Ouranopithecus* are accepted as homologous, many of the characters currently used to establish the pongine clade would seem to be primitive. The combination of characters in AS95-500 link together the European Middle and Late Miocene fossil apes in the genera *Dryopithecus*, *Ouranopithecus* and *Ankarapithecus* as stem members of the great ape and human clade and do not provide evidence for relationships with either the African apes²⁴ or pongines³⁰. □

Received 5 February; accepted 28 June 1996.

- Alpagut, B. & Martin, L. *Arastirma Sonuclari Toplantisi* **VIII**, 55–67 (1989).
- Alpagut, B., Fortelius, M. & Kappelman, J. *Arastirma Sonuclari Toplantisi* **XII**, 177–199 (1993).
- Kappelman, J. et al. in *Evolution of Western Eurasian Late Neogene Mammal Faunas* (eds Bernor, R., Fahlbusch, V. & Mittman, H.-W.) (Columbia Univ. Press, in the press).
- Andrews, P. & Tekkaya, I. *Palaeontology* **23**, 85–95 (1980).
- Ozansoy, F. *Bull. Miner. Res. Explor. Inst., Ankara* **49**, 29–48 (1957).
- Ozansoy, F. C. r. *hebd. Séanc. Acad. Sci., Paris* **240**, 992–994 (1955).
- Cande, S. C. & Kent, D. V. *J. geophys. Res.* **100**, 6093–6095 (1995).
- Bonis, de L., Bouvain, G., Geraads, D. & Koufos, G. *Nature* **345**, 712–714 (1990).
- Agusti, J. et al. *J. hum. Evol.* **31**, 143–155 (1996).
- Barry, J. C. in *Primate Evolution* (eds Else, J. G. & Lee, P. C.) 93–106 (Cambridge Univ. Press, Cambridge, 1986).
- Kappelman, J. et al. *J. hum. Evol.* **21**, 61–73 (1991).
- Simons, E. L. & Pilbeam, D. R. *Folia Primatol.* **3**, 81–152 (1965).
- Szalay, F. & Delson, E. *Evolutionary History of the Primates* (Academic, New York, 1979).
- Begun, D. R. & Gülec, E. *Am. J. phys. Anthropol. Suppl.* **20**, 63–64 (1995).
- Kelley, J. et al. *J. hum. Evol.* **28**, 503–517 (1995).
- Moya-Solà, S. & Köhler, M. *J. hum. Evol.* **29**, 101–139 (1995).
- Coolidge, H. J. Jr. *Am. J. phys. Anthropol.* **18**, 1–59 (1933).
- Cave, A. J. E. & Haines, R. W. *J. Anat.* **72**, 493–523 (1940).
- Ward, S. C. & Brown, B. in *Comparative Primate Biology Vol. 1: Systematics, Evolution, and Anatomy* (eds Swindler, D. R. & Erwin, J.) 413–452 (Liss, New York, 1986).
- Moya-Solà, S. & Köhler, M. *Nature* **365**, 543–545 (1993).
- Andrews, P. & Cronin, J. E. *Nature* **297**, 541–546 (1982).
- Ward, S. C. & Pilbeam, D. R. in *New Interpretations of Ape and Human Ancestry* (eds Ciochon, R. L. & Corruccini, R. S.) 211–238 (Plenum, New York, 1983).
- Bonis, de L. & Koufos, G. D. *J. hum. Evol.* **24**, 469–491 (1993).
- Begun, D. R. *J. hum. Evol.* **29**, 169–180 (1995).
- Martin, L. *Nature* **314**, 260–263 (1985).
- Begun, D. R. & Kordos, J. *hum. Evol.* **25**, 271–285 (1995).
- Bonis, de L. & Melentis, J. *Geobios* **10**, 849–885 (1977).
- Brown, B. & Ward, S. C. in *Orang Utan Biology* (ed. Schwartz, J. H.) 247–260 (Oxford Univ. Press, New York, 1988).
- Shea, B. T. *Am. J. phys. Anthropol.* **68**, 329–342 (1985).
- Moya-Solà, S. & Köhle, M. *Nature* **379**, 156–159 (1996).

ACKNOWLEDGEMENTS. The Sinap survey and salvage excavation projects have been conducted on behalf of the Republic of Turkey Ministry of Culture and directed by Ankara University and the Anatolian Museum of Civilization. We thank the General Directorate of Monuments and Museums, Ankara, for survey and excavation permits. We thank all members of the field crews especially Z. Bostan who found AS95-500, L. Martin, S. Sen, S. Baskan and S. Akay for providing the CT scans, and G. Conroy, M. Köhler, R. D. Martin, S. Moya-Solà and D. Pilbeam for comments on a draft of the manuscript. The Sinap project received support from the National Science Foundation, the Academy of Finland, the L. S. B. Leakey Foundation, Motorola, Inc., Autodesk, Inc., Ankara University, The University of Helsinki, The University of Texas at Austin, and the Museum of Anatolian Civilization.

CORRESPONDENCE and requests for materials should be addressed to J.K. (e-mail: jkappelman@mail.utexas.edu).

Infant leukaemia after *in utero* exposure to radiation from Chernobyl

E. Petridou*†, D. Trichopoulos*, N. Dessypris†, V. Flytzani†, S. Haidas‡, M. Kalmanti§, D. Kolioukas||, H. Kosmidis¶, F. Piperopoulou# & F. Tzortzatu*

* Harvard Center for Cancer Prevention, 677 Huntington Avenue, HSPH3-901 Boston, Massachusetts 02115, USA

† Department of Hygiene and Epidemiology, Athens University Medical School, Athens 115 27, Greece

‡ Department of Paediatric Haematology–Oncology, Agia Sophia Children's Hospital, Athens 115 27, Greece

§ Department of Paediatric Haematology–Oncology, University Hospital, Heraklion 715 00, Greece

|| Department of Paediatric Haematology–Oncology, Hippokrateion Hospital, Thessaloniki 546 42, Greece

¶ Department of Paediatric Haematology–Oncology, Aglaia Kyriakou Children's Hospital, Athens 115 27, Greece

Department of Paediatric Haematology–Oncology, Ahepa Hospital, Thessaloniki 546 36, Greece

* Oncology Unit, First Department of Paediatrics, University of Athens, Agia Sophia Children's Hospital, Athens 115 27, Greece

THERE has been no documented increase in childhood leukaemia following the Chernobyl accident. However, different forms of childhood leukaemia may not be equally susceptible to radiation carcinogenesis. Infant leukaemia is a distinct form associated with a specific genetic abnormality. Outside the former Soviet Union, contamination resulting from the Chernobyl accident has been highest in Greece and Austria and high also in the Scandinavian countries^{1–4}. All childhood leukaemia cases diagnosed throughout Greece since 1 January 1980 have been recorded. Here we report that infants exposed *in utero* to ionizing radiation from the Chernobyl accident had 2.6 times the incidence of leukaemia compared to unexposed children (95% confidence interval, 1.4 to 5.1; $P \approx 0.003$), and those born to mothers residing in regions with high radioactive fallout were at higher risk of developing infant leukaemia. No significant difference in leukaemia incidence was found among children aged 12 to 47 months. Preconceptional irradiation had no demonstrable effect on leukaemia risk at any of the studied age groups.

The Chernobyl accident, which occurred on 26 April 1986, has led to widespread radionuclide contamination. In Belarus and Ukraine an increase in the incidence of thyroid cancer has been reported^{5,6}, but no increase in childhood leukaemia has been documented so far either in these^{6,7} or any other European country^{8,9}, including Greece⁴. However, different forms of childhood leukaemia may differ in susceptibility to radiation carcinogenesis. Evidence has emerged that infant leukaemia is a distinct entity associated, in more than two-thirds of cases, with a specific genetic abnormality in the 11q23 chromosome band^{10,11}. Moreover, the relevant exposure period for infant leukaemia is likely to be during the intrauterine life, when susceptibility to the carcinogenic effects of ionizing radiation is assumed to be particularly high^{12,13}.

All childhood leukaemia cases diagnosed throughout Greece from 1980 have been recorded by a national network of childhood oncologists^{4,14}. At the time of our analysis, the records were complete to 31 December 1994. For every leukaemia case, gender, date of birth, date of diagnosis, type of leukaemia, and maternal residence are available. Fallout radiation from the Chernobyl accident has been measured by other investigators in more than 1,000 samples of surface soil from 42 of the 52 administrative divisions of Greece; the remaining 10 administrative divisions were exposed to a lesser extent for meteorological reasons (measurements summarized in ref. 4). Fallout radio-

activity from Chernobyl was estimated by taking the arithmetic mean of the measurements in every division and is expressed in Bq kg⁻¹ of ¹³⁷Cs in surface soil. The administrative divisions were then combined into 3 categories, with fallout radioactivity 1,000 or more Bq kg⁻¹; 999–100 Bq kg⁻¹; and below 100 Bq kg⁻¹.

Exposure of the Greek population to Chernobyl radiation started soon after the accident and was notable for about a year; average exposure has been estimated at about 2 mSv (ref. 1). Therefore, children born during the second half of 1986, the first half of 1987, and most of those born during the second half of 1987 were considered as exposed to Chernobyl irradiation *in utero*, whereas those born from 1 January 1980 to 31 December 1985 (6 years) and those born from 1 January 1988 to 31 December 1990 (3 years) were considered unexposed (Table 1). We have not classified children born during the first half of 1986, because those born before the end of April were clearly unexposed, whereas those born in May and June were only exposed during the last pregnancy months, when the risk of radiation-induced carcinogenesis is unknown.

Our analysis was based on estimation of age-adjusted leukaemia rates (as in ref. 15), calculated until the end of the fourth year of life, because data beyond that age were not available for the children born during 1990. The results are summarized in Table 1. Comparison of leukaemia rates between the two unexposed birth cohorts (born before 31/12/85 or after 1/1/88), provides no indication of any difference in incidence (for infants, $X^2 = 0.39$, $P = 0.53$; for children 12 to 47 months, $X^2 = 0.02$, $P = 0.90$). There is also no difference in the incidence of leukaemia at ages from 12 to 47 months between children exposed and unexposed *in utero* to Chernobyl radiation ($P \approx 0.56$). In contrast, *in utero*-exposed infants have 2.6 times the incidence of leukaemia compared to nonexposed infants ($P \approx 0.003$). Changing the operational definition of exposure by 6 months had no discernible effect.

The incidence of infant leukaemia among unexposed children was 27.9 per 10⁶ person-years (95% confidence interval, CI, 18.9–39.5). Among children exposed *in utero* the corresponding incidence rates were 32.2 (CI 1.6–159.8) for those born in low radioactivity administrative divisions (only 1 case); 71.4 (CI 31.2–141.3) for those born in intermediate radioactivity divisions (7 cases); and 116.6 (CI 37.0 to 281.3) for those born in high radioactivity divisions (4 cases). The latter two incidence rates are significantly higher than that among children who were unexposed *in utero* (two tailed P values 0.02 and 0.004, respectively; two tailed P for trend 0.0005).

Several molecular rearrangements in 11q 23 (the MLL gene) are far more common in infant leukaemia than in leukaemia of older children or adults^{16–19}, and they are likely to originate from mutations during pregnancy²⁰. Ionizing radiation is an established cause of acute lymphoblastic leukaemia and the intrauterine life represents a period of increased susceptibility^{12,21}. As infant leukaemia is likely to have prenatal origin of non-constitutive nature^{19,20,22} ionizing radiation *in utero* becomes a highly plausible cause of this disease.

Previous studies have not specifically examined infant leukaemia, but in a Swedish study²³ three infants that were *in utero* at the time of the accident developed leukaemia. Studies evaluating the effect on childhood cancer of prenatal exposure to diagnostic X-ray examinations did not indicate that infant leukaemia is differentially susceptible to radiation carcinogenesis^{21,24,25}. However, prenatal diagnostic X-rays are almost always undertaken during the last months of pregnancy and in discrete sessions. It is possible that the early pregnancy period, the time of exposure to radiation in our study, represents a high-risk stage or contains a window of increased susceptibility. Data concerning dated X-rays in pregnancy at different stages of intrauterine development are remarkably consistent with this view²⁵, although they do not specifically indicate that infant leukaemia is differentially susceptible to early fetal irradiation.

Our study also provides evidence against the hypothesis²⁶ that paternal exposure to ionizing radiation increases the risk of either

TABLE 1 Children of specified birth cohorts who developed leukaemia in Greece

Birth cohort	Liveborn	Characterization	Number of cases grouped by age at diagnosis (months)*				
			–5	6–11	12–23	24–35	36–47
1/1/80 to 31/12/85	801,175	unexposed	8	14	55	64	68
1/7/86 to 31/12/87	163,337	exposed	4	8	8	19	16
1/1/88 to 31/12/90	311,391	unexposed	3	6	17	36	26
Incidence rate (per 10 ⁶ person-years)		exposed	49.0	98.0	49.0	116.3	98.0
		unexposed	19.8	36.0	64.7	89.9	84.5
Rate ratio			2.6			1.1	
95% confidence interval			1.4–5.1			0.8–1.5	
P (2-tailed)			P ≈ 0.003			P ≈ 0.56	

* Follow-up is 0.5 year in the first 2 columns, 1 year in the last 3.

infant leukaemia or leukaemia at ages 12 to 47 months, and supports the counter-arguments advanced in this context^{27,28}. Thus, leukaemia rates were almost identical among children born between 1/1/80 and 31/12/85 and those born between 1/1/88 and 31/12/90.

In conclusion, we provide evidence that infant leukaemia may be caused by very low level intrauterine exposure to ionizing radiation; that fallout from the Chernobyl explosion may have increased the incidence of infant leukaemia among Greek children exposed *in utero*, perhaps by as much as 2 to 3 fold; and that low-level preconceptional radiation has no demonstrable effect on leukaemia risk. □

Received 4 April; accepted 3 June 1996.

1. Papastephanou, C., Manolopoulou, M. & Charalambous, S. J. *Envir. Radioact.* **7**, 49–64 (1988).
2. Auvinen, A. et al. *Br. med. J.* **309**, 151–154 (1994).
3. International Chernobyl Project Technical Rep. (IAEA report number: ISBN 92-0-129191-4, Vienna, 1992).
4. Petridou, E. et al. *Scand. J. Soc. Med.* **22**, 127–131 (1994).
5. Abelin, T. et al. *Soz. Präventivmed.* **39**, 189–197 (1994).
6. Kuchuk, A. A. *Tox. Lett.* **72**, 213–217 (1994).
7. Vanchieri, C. J. *natn. Cancer Inst.* **84**, 1616 (1992).
8. Hjalmar, U., Kulldorff, M. & Gustafsson, G. *Br. med. J.* **309**, 154–157 (1994).

9. Parkin, D. M. et al. *Euro. J. Cancer* **29**, A87–95 (1992).
10. Cirino, G. *Blood* **82**, 544–546 (1993).
11. Cortes, J. E. & Kantarjian, H. M. *Cancer* **76**, 2393–2416 (1995).
12. U.S. National Academy of Sciences Committee on Biological Effects of Ionizing Radiation. *BEIR-V Rep.* (U.S. N.A.S. Washington, 1990).
13. Hoshimoto, Y., Kato, H. & Schull, W. J. *RERF Technical Report 4–88* (Radiation Effects Research Foundation, Hiroshima, 1988).
14. Petridou, E. et al. *Br. J. Cancer* (in the press).
15. Rothman, K. J. & Boice, J. D. Jr Publication 79–1649 (National Institutes of Health, 1979).
16. Ross, J. A., Davies, S. M., Potter, J. D. & Robison, L. L. *Epidemiol. Rev.* **16**, 243–271 (1994).
17. Clearly, M. L. *New Engl. J. Med.* **329**, 958–960 (1993).
18. Canaani, E., Nowell, P. C. & Croce, C. M. *Adv. Cancer Res.* **66**, 213–234 (1995).
19. Pui, C.-H., Kane, J. R. & Crist, W. M. *Leukemia* **9**, 762–769 (1995).
20. Ford, A. M. et al. *Nature* **363**, 358–360 (1993).
21. Monson, R. R. & MacMahon, B. in *Radiation Carcinogenesis: Epidemiology and Biological Significance* (eds Boice, J. D. & Fraumeni, J. F. Jr) 97–105 (Raven, New York, 1984).
22. MacMahon, B. & Levy, M. A. *New Engl. J. Med.* **270**, 1082–1085 (1964).
23. Ericson, A. & Kallen, B. *Envir. Res.* **67**, 149–159 (1994).
24. Bithell, J. F. & Stewart, A. M. *Brit. J. Cancer* **31**, 271–287 (1975).
25. Mole, R. H. *Brit. J. Cancer* **62**, 152–168 (1990).
26. Gardner, M. J. et al. *Br. med. J.* **300**, 423–429 (1990).
27. MacMahon, B. *Cancer Causes Contr.* **3**, 283–288 (1992).
28. Doll, R., Evens, H. J. & Darby, S. C. *Nature* **367**, 678–680 (1994).

ACKNOWLEDGEMENTS. We thank B. MacMahon for his constructive presubmission review of this paper. This study was supported in Greece by the 'Europe Against Cancer' programme of the European Union, and in Boston by a grant to Harvard University from G. S. Livanos.

CORRESPONDENCE and requests for materials should be addressed to D.T. (e-mail: epetrid@atlas.uoa.ariadne-t.gr).

Correct *Hox* gene expression established independently of position in *Caenorhabditis elegans*

Deborah Cowing & Cynthia Kenyon

Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, California 94143-0554, USA

THE *Hox* genes are expressed in a conserved sequence of spatial domains along the anteroposterior (A/P) body axes of many organisms¹. In *Drosophila*, position-specific signals located along the A/P axis establish the pattern of *Hox* gene expression^{2–4}. In the nematode *Caenorhabditis elegans*, it is not known how the pattern of *Hox* gene expression is established. *C. elegans* uses lineal control mechanisms and local cell interactions to specify early blastomere identities^{5,6}. However, many cells expressing the same *Hox* gene are unrelated by lineage, suggesting that, as in *Drosophila*, domains of *Hox* gene expression may be defined by cell-extrinsic A/P positional signals. To test this, we have investigated whether posterior mesodermal and ectodermal cells will express their normal posterior *Hox* gene when they are

mispositioned in the anterior. Surprisingly, we find that correct *Hox* gene expression does not depend on cell position, but is highly correlated with cell lineage. Thus, although the most striking feature of *Hox* gene expression is its positional specificity, in *C. elegans* the pattern is achieved, at least in part, by a lineage-specific control system that operates without regard to A/P position.

In *C. elegans*, homologues of the *Drosophila* genes *Sex combs reduced*, *Antennapedia* and *Abdominal-B* are expressed in consecutive domains along the A/P body axis^{7,8}. Cells that express a particular *Hox* gene have only position in common; they are unrelated by cell type or lineal origin (Fig. 1). To begin to understand how the *Hox* gene expression pattern is established, we have mispositioned cells that are normally located in the posterior body region, and have investigated whether they would still express the posterior-specific *Hox* gene, the *Antennapedia* homologue *mab-5*.

A mesodermal cell called M is unusual in that it migrates a long distance from its anterior origin into the posterior body region⁹ (Fig. 2a). As it reaches the posterior, it begins to express *mab-5* (Fig. 2c–e). In contrast, cells closely related to M by lineage remain in the anterior and do not express *mab-5* (ref. 9, and data not shown). Thus it seemed likely that M switches on *mab-5* expression because it encounters a signal located in the posterior body region. To change the position of this cell, we blocked its migration by using cytochalasin D, which blocks actin polymerization, or colchicine, which depolymerizes microtubules^{10,11}. If M

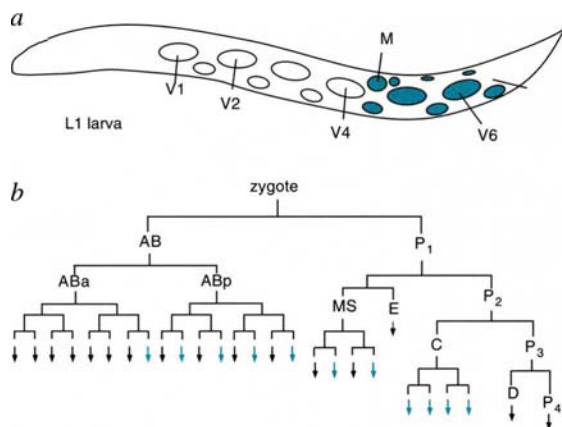


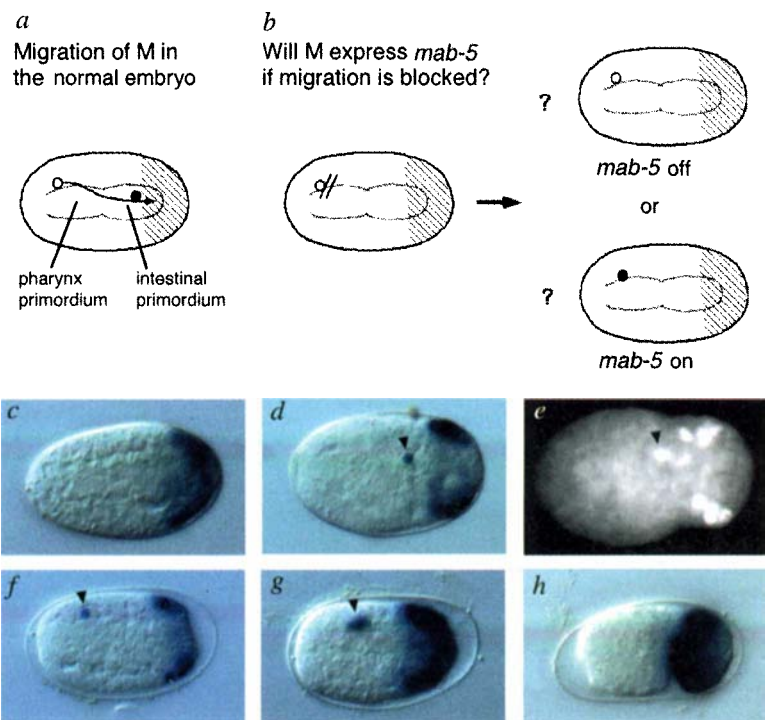
FIG. 1 The locations and lineal relationships of *mab-5*-expressing cells. *a*, Cells expressing *mab-5* are located in the posterior of the animal. L1 larva, with representative cells that express, or whose progeny express, *mab-5* are shown in blue⁷. The locations of M, V1, V2, V4 and V6 are indicated. *b*, Cells expressing *mab-5* arise from many branches of the lineage. These cells are positioned adjacent to one another in the posterior as a result of their prior cleavage orientation⁹. In this truncated representation of the *C. elegans* lineage, branches giving rise to one or more *mab-5*-expressing cells are indicated in blue; each of these branches produces many cells that do not express *mab-5*.

requires signals located in the posterior body region to turn on *mab-5*, we would not expect it to express *mab-5* in the anterior (Fig. 2*b*). We found that these drugs did not affect the overall pattern of *mab-5* expression, indicating that they did not cause a nonspecific disruption in patterning information (Fig. 2). How-

ever, surprisingly, we now saw embryos in which M was located far in the anterior but nevertheless expressed *mab-5-lacZ* (Fig. 2*f-h*). To investigate whether the displaced M cells might extend processes into the posterior body region that could potentially sense a positional signal, we repeated the experiment using a cytoplasmically localized *mab-5-lacZ* reporter. We did not observe any processes extending from the displaced M cells (Fig. 2*g*). Thus we conclude that the ability of M to express *mab-5* does not depend on its position in the posterior.

We next investigated whether the ability of other cells to express *mab-5* was correlated with their position in the embryo, or with their cell lineage. To do this, we made use of the finding that cell interactions at the four-cell stage are required to specify blastomere identities^{12,13}. When the signalling blastomeres EMS and P₂ are killed, or when the signalling system is inactivated by mutation, the two remaining blastomeres, ABa and ABp, generate incorrect lineages: together, they generate four exact copies of the lineage normally produced by the blastomere ABap, one of the grand-daughters of ABa (Fig. 3*a*). During normal development, only two descendants of ABap, the lateral epidermal cells V6L (left) and V6R (right), express *mab-5*. All the remaining cells generated by ABap assume positions in the head or anterior body regions, and do not express *mab-5* (ref. 9). Following these ablations, eight cells are predicted to be lineally equivalent to V6L or V6R. Therefore, if events correlated with cell lineage influence *mab-5* expression, it seemed possible that these eight cells, referred to here as 'V6 equivalents', would express *mab-5* regardless of their position in the embryo. Alternatively, if cell-extrinsic A/P positional information alone regulates *mab-5* expression, we would expect a different outcome. Expression of *mab-5* might still be restricted to cells located in the posterior, or, if ablation of EMS and P₂ destroys the hypothetical positional information, no cells should express *mab-5*. Finally, if the ablation

FIG. 2 Expression of *mab-5* in mislocalized M cells. *a*, M originates in the anterior and migrates posteriorly⁹. Shading indicates *mab-5* expression in other posterior cells. (Note that *mab-5* activity is not required for M migration¹⁷ or for *mab-5* expression.) Open circle, birthplace of M; filled circle, position where *mab-5* expression is first seen in M. *b*, Experimental design. We blocked the migration of M at or soon after its birth, and investigated whether M could still express *mulS3*, an integrated *mab-5-lacZ* fusion¹⁸. *c*, Expression of *mab-5-lacZ* in untreated embryos soon after the birth of M, 3.5 h after first cleavage (25 °C), and *d*, at 4.5 h. Arrowhead indicates M. To determine when *mab-5-lacZ* was first expressed in M, staged embryos carrying *mab-5-lacZ* (*mulS3*) were fixed and stained¹⁹. M expressed *mab-5-lacZ* in 0 of 46 embryos at 3.5 h, 4 of 61 at 4 h, and 11 of 34 4.5 h after first cleavage (25 °C). *e*, Expression of *mab-5* visualized by antibody staining. Mixed-stage embryos were fixed and stained (procedure modified from ref. 18). Antibody and *lacZ* staining patterns were similar in embryos and larva (not shown). In *d* and *e*, M has migrated to just anterior of the two germline precursors, Z2 and Z3. *f*, *g*, Embryos treated with cytochalasin D during M migration. *f*, The nuclear-localized *mab-5-lacZ* fusion *mulS3* was used; *g*, the cytoplasmic fusion *mulS12*, which stains neuronal processes, was used to visualize any M cell process. *h*, Untreated control. Cytochalasin-treated embryos are unable to elongate²⁰; control embryos elongate to comma stage during this period. By this time, M should have reached the posterior body region, which contains many other stained cells. Embryos staged early during the migration of M were permeabilized with a laser microbeam in culture medium²¹ containing no drug, cytochalasin D (sigma C-8273, 16 µM) or colchicine (Sigma C-9754, 1 mM). Embryos were incubated for 2 h at 25 °C to allow *mab-5-lacZ* expression, then fixed and stained¹⁹. Overall, the *mab-5* expression pattern appeared normal. However, we now observed anteriorly located cells expressing *mab-5-lacZ* that could be identified as M by their location in the birthplace of M, just left of the pharynx, or along the normal migration path of M. In 12 of 116 cytochalasin-treated embryos (10%), and 1 of 11 colchicine-treated embryos (9%), the stained M cell was further anterior than was ever



seen in untreated embryos (compare *d* and *f*; anterior is to the left). In 41 of 116 cytochalasin-treated embryos (35%), and 5 of 11 colchicine-treated embryos (45%), M was found in the central region, still anterior to the remainder of *mab-5*-expressing cells. M was not found in the anterior or central region in control embryos (0 of 87).

has the effect of redistributing cell-extrinsic positional information, then cells in both the posterior and anterior might express *mab-5*, but there should not be a correlation between cell lineage and *mab-5* expression.

We killed the signalling blastomeres EMS and P_2 in embryos carrying a *mab-5-gfp* reporter fusion¹⁴, and recorded the development of the embryo using time-lapse video microscopy¹⁵. We then visualized *mab-5-gfp* expression using epifluorescence. To determine whether the ablations had elicited the expected lineage transformations, we observed the cell divisions leading to the predicted V6 equivalents. In some cases, the cells of interest were obscured by overlying cells, so it was not possible to follow their development. In addition, in some embryos, the lineage transformations appeared to be incomplete (see Fig. 4). However, we were able to identify a total of 15 unambiguous V6 equivalents from 6 embryos (see Fig. 4). In 11 of 15 cases, these cells expressed *mab-5* (Figs 3 and 4). Because expression of the *mab-5-gfp* fusion was not fully penetrant even in normal embryos (data not shown), we conclude that V6 equivalents often had the potential to express *mab-5*.

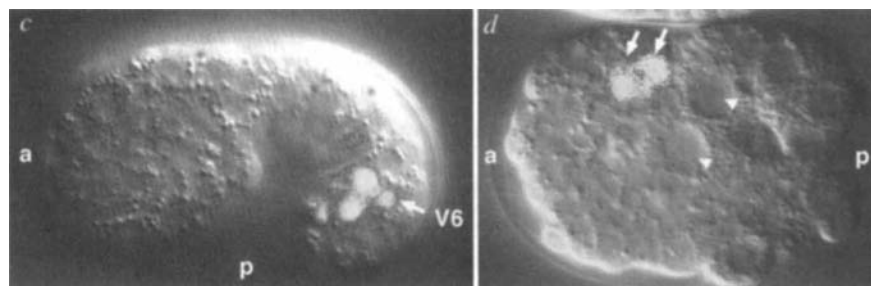
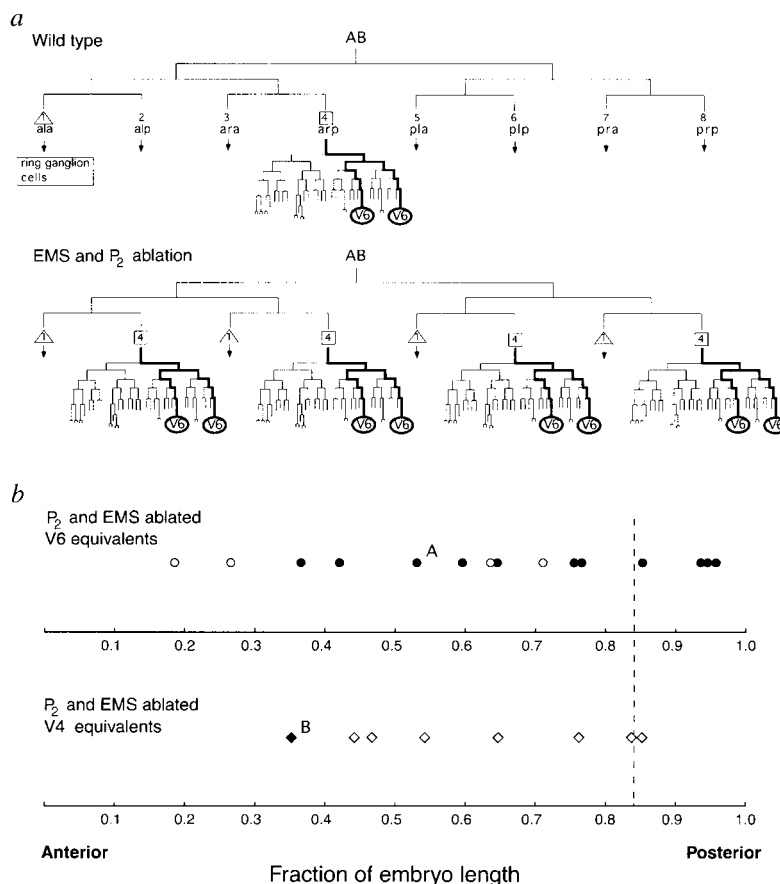
Next we investigated whether *mab-5* expression was restricted

to the V6 equivalents. In these embryos, few cells expressed *mab-5* (Fig. 4), as would be expected from the lineage transformations if this were the case (see Fig. 3a). Altogether, 13 of the *mab-5*-expressing cells could be assigned to specific positions in the embryonic lineages, of which all but two were found to be V6 equivalents; one exception was a lineal equivalent of the head epidermal cell H2, and the other is described below. As a negative control, we also followed the lineages of the cousins (V1, V2) and sisters (V4) of the V6 equivalents (Fig. 3b and 4). During normal development, V1, V2 and V4 become lateral epidermal cells, as does V6. These cells are located anterior to V6, and do not express *mab-5*. We found that none of the V1 or V2 equivalents expressed *mab-5* (Fig. 4). Of eight V4 equivalents, one expressed *mab-5* (Figs 3 and 4); however, the development of this cell was unusual (see Fig. 3 legend). Together these observations indicate that cells that are lineally equivalent to *mab-5*-expressing cells are likely to express *mab-5*; in contrast, cells that are not lineally equivalent to *mab-5*-expressing cells are not. Significantly, we observed that *mab-5-gfp* expression was not correlated with position in these embryos. V6 equivalents expressing *mab-5* could be found in either posterior or anterior positions (Fig. 3b, c); moreover, V6

FIG. 3 Anteriorly located V6 equivalents can express *mab-5*.

a, Generation of V6 equivalents after P_2 and EMS ablation. The descendants of AB at the AB₈ cell stage are numbered (numbering scheme modified from ref. 13). The embryonic lineage of AB₈ is shown, with the divisions giving rise to the two V6 cells, V6L and V6R, shown in bold. V4(L/R) are the sisters of V6(L/R), respectively. After killing EMS and P_2 (or in *gfp-1* embryos), lineages 1 and 4 are reiterated^{12,13,22}. Lineage 1 gives rise to neurons and support cells in the head⁹. During normal development, none of these cells expresses *mab-5*. **b**, A/P positions of V6 and V4 equivalents. Top, V6 equivalents (circles); bottom, V4 equivalents (diamonds). Filled symbols, cells expressing *mab-5-gfp*; open symbols, no expression of *mab-5-gfp*. The broken line, position of V6 in control embryos; *gfp-1* embryos were also assayed, and 5 of 5 V6 equivalents expressed *mab-5-gfp* and were located in the posterior (not shown). The cell labelled A was identified in an experiment in which only EMS was ablated. The behaviour of the one GFP-expressing V4 equivalent (labelled B) was surprising. Normally V4 remains anterior to its sister V6, and does not express *mab-5*. In this lineage, the two sister cells rotated about one another, with the V4 equivalent ending up posterior to the V6 equivalent, and thus in the normal V6 position. When later examined for GFP expression, the V4 equivalent expressed GFP, but the V6 equivalent did not. Because of their rotation, however, the pair looked similar to other V4/V6 pairs in that the more posterior cell (normally V6) expressed *mab-5-gfp*. This was unexpected, and cannot be explained by a simple lineal or positional model, as both sister cells were located in the anterior. It is not known whether this rotation ever occurs during normal development. **c**, **d**, Expression of *mab-5-gfp* in control embryo and after P_2 and EMS ablation (overlay of the fluorescent and Nomarski images). **c**, Control, unablated, comma-stage embryos; arrow indicates V6. **d**, Embryo in which P_2 and EMS were ablated, 5.3 h after the ablations (at 25 °C), the equivalent of comma stage; arrows indicate two V6 equivalents. Large cells (arrowheads) are descendants of P_2 and EMS. In these embryos EMS and P_2 divided a few times, albeit very late and abnormally.

METHODS. P_2 and EMS were ablated in 4-cell embryos from a strain carrying *mulS16*, a *mab-5-gfp* fusion containing 10 kb upstream of *mab-5*. P_2 was irradiated for 30 s immediately after cytokinesis; EMS was irradiated for 1 min following the ablation of P_2 . Embryonic development was recorded on a 4-D time-lapse system¹⁵ for 5 h at 25 °C. At the end of the recording period, illumination was switched from Nomarski to epifluorescence, and GFP-expressing cells were recorded.



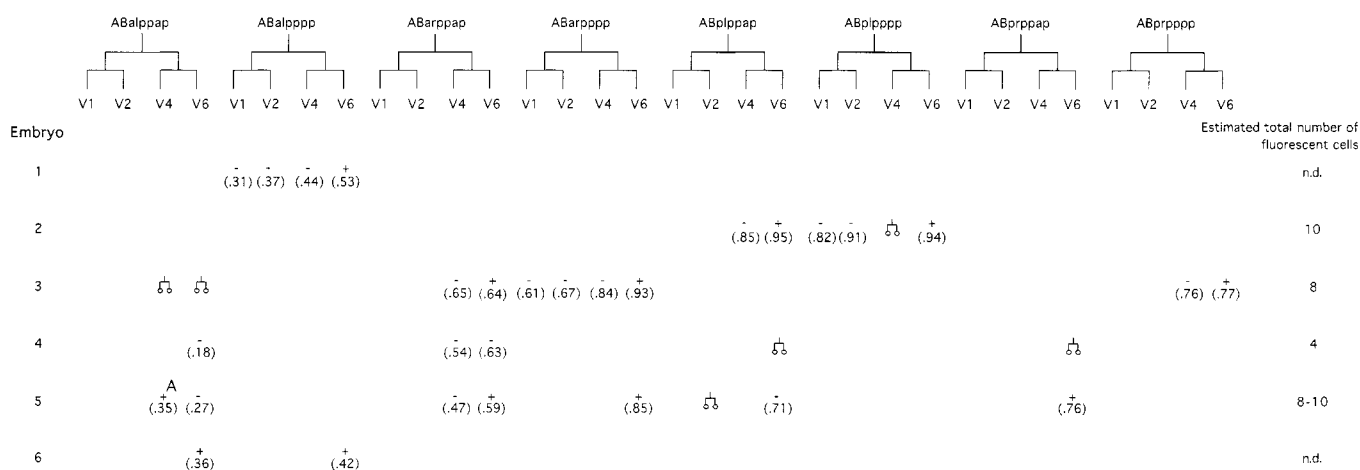


FIG. 4 Expression of *mab-5-gfp* in V-cell equivalents. Data are shown for all V-cell lineages that could be reconstructed from the video recordings of the 6 experimental embryos. Each of the 8 lineages that would produce V6 equivalents in a single fully transformed animal are shown across the top (see Fig. 3a). +, *mab-5-gfp* expression; –, no expression. The position of the cell along the A/P axis (as in Fig. 3b) is indicated in parentheses. In embryo 1, only EMS was ablated. This causes transformation of only ABalp to the ABarp fate²³ (see Fig. 3a), thus only ABalp and ABarp descendants were examined. In embryo 6, ABp was inadvertently briefly exposed to the laser beam; to avoid analysing the effects of possible laser damage, only ABa descendants were examined. In some cases cells that were expected to have a V-cell identity instead divided (lineage symbol). This is the fate expected for these cells in a control, unablated embryo⁹, suggesting that

these lineage transformations were incomplete; these were not considered further. In total, 0 of 3 V1, 0 of 3 V2, 1 of 8 V4 and 11 of 15 V6 equivalents expressed *mab-5-gfp*. For embryos 2–5, the total number of fluorescent cells was estimated. Because in lower focal planes it was often difficult to determine whether an area of fluorescence represented one or more cells, these numbers are imprecise. However, the eight V6 equivalents are the only *mab-5*-expressing cells expected if expression is correlated with lineage history; the estimated numbers of fluorescent cells are close to this, taking into account the incomplete penetrance of GFP expression and incomplete lineage transformations we observed. Furthermore, only two expressing cells whose lineage histories could be determined were not V6 equivalents (see text.) The *mab-5*-expressing V4 equivalent labelled A behaved abnormally (see legend to Fig. 3).

equivalents expressing *mab-5-gfp* could be found anterior to non-expressing V1, V2 and V4 equivalents in the same embryo (Figs 3 and 4).

In summary, we found that two different types of cells, the mesodermal cell M and the ectodermal V6 equivalents, are capable of expressing *mab-5* even if they are located in the incorrect body regions. These findings imply that mechanisms other than cell-extrinsic A/P positional signals play an important role in regulating expression of this *Hox* gene in *C. elegans*. However, our findings do not exclude the possibility that A/P positional information regulates *Hox* gene expression in some cells, or that it acts redundantly with other mechanisms to ensure that *Hox* gene expression occurs in a highly reproducible fashion.

The strong correlation between cell lineage and *mab-5* expression suggests that cells may inherit *mab-5*-regulatory information from their ancestors. This information could be segregated asymmetrically to the precursors of *mab-5*-expressing cells during cell division, or it could be imparted to these cells by intercellular signalling, or both. The results of our blastomere-ablation experiments indicate that intercellular signals are not sufficient to control *mab-5* expression, however. If this were the case, then

any hypothetical signalling cells that were still present in embryos in which P₂ and EMS had been ablated should have induced *mab-5* expression in small clusters of surrounding cells, regardless of their lineages. Instead, the vast majority of *mab-5*-expressing cells were lineally equivalent to V6. Thus if such signalling cells do exist, then remarkably tight lineage-specific controls must determine which cells can respond.

The correlation between cell lineage and ability to express *mab-5* was particularly surprising because the posterior cells that express *mab-5* during normal development are not related by lineage. Thus a very complex regulatory system may be required to establish what looks like a simple pattern of gene expression. Together with the previous finding that positional specificity in the leech is achieved by temporal rather than spatial control mechanisms¹⁶, our results suggest that the mechanisms by which different organisms regulate expression of the conserved *Hox* genes have diverged substantially during evolution. It will be interesting to learn how position-independent mechanisms for patterning *Hox* gene expression operate at the molecular level. This information should provide clues to how this complex way of regulating *Hox* gene expression could have arisen during evolution. □

Received 18 March; accepted 30 May 1996.

- McGinnis, W. & Krumlauf, R. *Cell* **68**, 283–302 (1992).
- Lawrence, P. A. & Morata, G. *Cell* **78**, 181–190 (1994).
- McGinnis, W. et al. *Adv. Genet.* **27**, 363–402 (1990).
- Riddle, R. D., Johnson, R. L., Laufer, E. & Tabin, C. *Cell* **75**, 1401–1416 (1993).
- Priess, J. R. *Curr. Biol.* **4**, 563–568 (1994).
- Wood, W. B. & Edgar, L. G. *Trends Genet.* **10**, 49–54 (1994).
- Wang, B. B. et al. *Cell* **74**, 29–42 (1993).
- Clark, S. G., Chisholm, A. D. & Horvitz, H. R. *Cell* **74**, 43–55 (1993).
- Sulston, J., Schierenberg, E., White, J. & Thomson, J. *Dev. Biol.* **100**, 64–119 (1983).
- Singer, S. J. & Kupfer, A. A. *Rev. Cell Biol.* **2**, 337–365 (1986).
- Dustin, P. *Microtubules* (Springer, New York, 1984).
- Hutter, H. & Schnabel, R. *Development* **120**, 2051–2064 (1994).
- Moskowitz, I. P. G., Gendreau, S. B. & Rothman, J. H. *Development* **120**, 3325–3338 (1994).
- Chalfie, M., Tu, Y., Eusirchen, G., Ward, W. W. & Prasher, D. C. *Science* **263**, (1994).
- Hird, S. N. & White, J. G. *J. Cell Biol.* **121**, 1343–1355 (1993).

- Nardelli-Haeffiger, D., Bruce, A. E. & Shankland, M. *Development* **120**, 1839–1849 (1994).
- Kenyon, C. *Cell* **46**, 477–487 (1986).
- Salsar, S., Loer, C. & Kenyon, C. *Genes Dev.* **7**, 1714–1724 (1993).
- Cowing, D. W. & Kenyon, C. *Development* **116**, 481–490 (1992).
- Priess, J. & Hirsch, D. *Dev. Biol.* **117**, 156–173 (1986).
- Cowan, A. & McIntosh, J. *Cell* **41**, 923–932 (1985).
- Mello, C. C., Draper, B. W. & Priess, J. R. *Cell* **77**, 95–106 (1994).
- Hutter, H. & Schnabel, R. *Development* **121**, 1559–1568 (1995).

ACKNOWLEDGEMENTS. C. Hunter constructed, integrated and kindly provided the *mab-5-gfp* fusion. We thank L. Honigberg and M. Mueller for constructs and strains; S. Salsar for *Mab-5* antibody and advice on antibody staining; I. Zipkin for observations on M migration behaviour; and members of our laboratory for ideas, discussions and comments on the manuscript. This work was supported by an NIH grant to C.K., who was a Packard Foundation Fellow during this time.

CORRESPONDENCE and requests for materials should be addressed to C.K. (e-mail: cynthia_kenyon.biochem@quickmail.ucsf.edu).

BMP-4-responsive regulation of dorsal–ventral patterning by the homeobox protein Mix.1

Paul E. Mead^{*†}, Iman H. Brivanlou^{*‡}, Clair M. Kelley^{*} & Leonard I. Zon^{*†}

^{*} Division of Hematology/Oncology and [†] Howard Hughes Medical Institute, Children's Hospital, Harvard Medical School, 300 Longwood Avenue, Enders 650, Boston, Massachusetts 02115, USA

[‡] Present address: Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts 02138, USA.

In an expression screen for factors that pattern or induce ventral mesoderm, we isolated a complementary DNA encoding Mix.1, a paired class homeobox gene with no previously known function¹. Injection of Mix.1 messenger RNA results in extensive blood formation in the whole embryo and transforms dorsal mesoderm to a ventral fate. Mix.1 expression is induced by bone morphogenetic protein-4 (BMP-4), and a dominant inhibitory mutant of Mix.1 can restore a dorsal axis in embryos ventralized by ectopic BMP-4 expression. Mix.1 can form heterodimers with the dorsalizing gene *siamois*, which encodes a homeodomain protein that is structurally similar to Mix.1. Furthermore, Mix.1 blocks the duplicated axis induced by ectopic *siamois* expression. Our findings indicate that Mix.1 participates in a BMP-4 signalling pathway to pattern ventral mesoderm, and suggest a model whereby dimerization of homeodomain proteins regulates dorsal–ventral patterning.

To isolate molecules involved in ventral mesoderm formation, we used a modified expression cloning strategy in *Xenopus* embryos². An expression library was generated from ultraviolet-ventralized³ embryos collected at late blastula to early gastrula stages. Synthetic mRNA was prepared from each pool of 5,000 clones and injected into fertilized embryos before the first cell cleavage. Two days after injection, embryos were scored for ventral phenotypes which are characterized by loss of head structures and increased blood formation (Fig. 1a). Seven of the ninety-two cDNA pools produced ventral phenotypes, and one pool was sib-selected by sequential fractionation to smaller sized

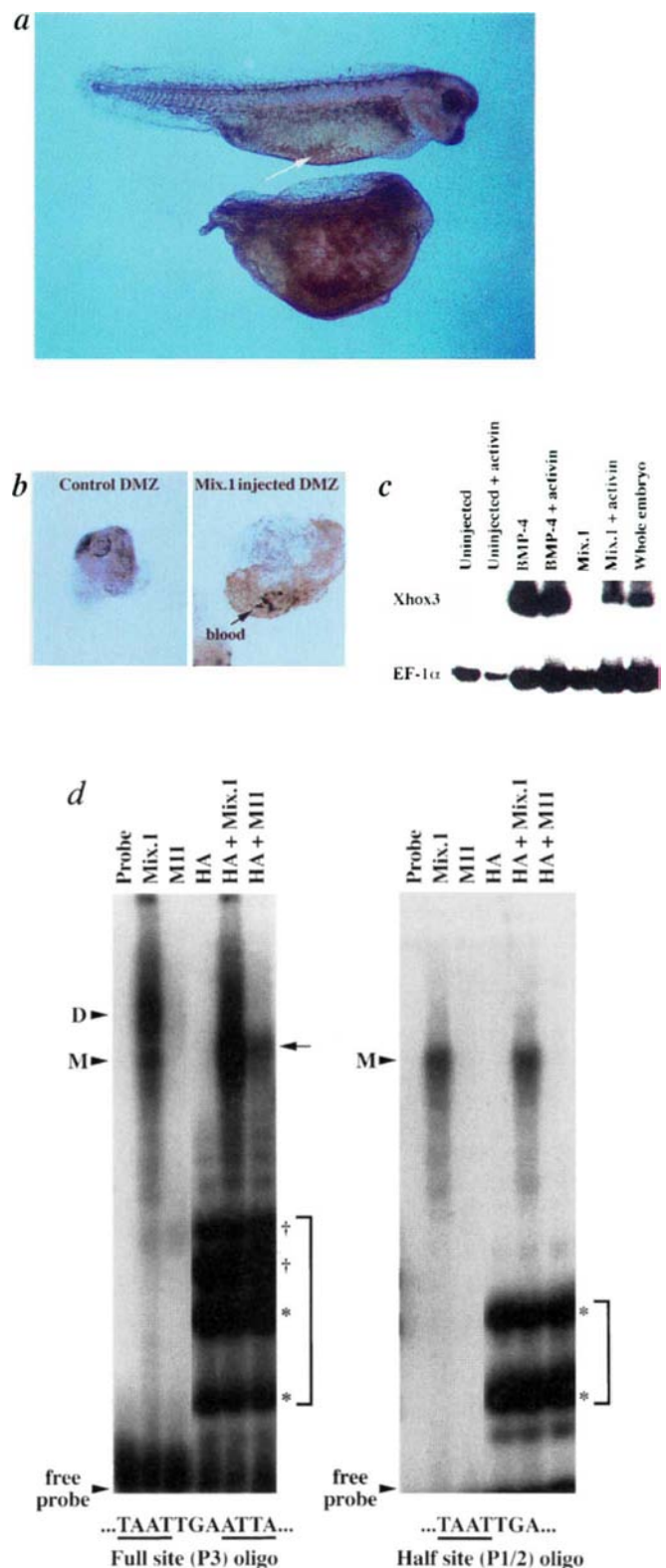


FIG. 1 Effect of Mix.1 RNA injection on whole embryos and embryonic explants. **a**, The top embryo is an uninjected control demonstrating globin staining (brown) in the ventral blood island (white arrow). Below is a sibling embryo injected with 1 ng Mix.1 RNA at the one-cell stage. The ventralized phenotype is characterized by a complete loss of head structures and globin expression throughout the embryo. **b**, Mix.1 transforms dorsal mesoderm to a ventral fate. DMZ is dissected from an embryo injected with Mix.1 RNA (250 pg Mix.1 per blastomere at the two-cell stage), cultured to control-stage 36 and immunostained for larval globin (arrow). Control (β -galactosidase-injected) DMZ explants do not express globin. **c**, Mix.1 does not induce mesoderm, but patterns existing mesoderm to a ventral fate. Injection of Mix.1 RNA alone does not induce the expression of ventral posterior marker *Xhox3* in animal pole cells. However, Mix.1 is capable of directing an activin-treated animal cap to express *Xhox3*. **d**, Gel-mobility-shift analysis²⁰ of *in vitro* translated Mix.1. Mix.1 monomers (M) were detected using the half-site probe (P_{1/2}); monomer and dimer complexes (M and D, respectively) were identified using the tandem repeat full-site probe (P₃)¹³. The dominant inhibitory mutant M11 binds DNA weakly compared to wild-type Mix.1. A similar mutation in Antennapedia class homeodomains has been shown to disrupt DNA binding²¹. A deletion mutant containing the homeodomain alone (HA) was mixed with full-length proteins to identify heterodimers. The *in vitro* translated homeodomain protein contained a degradation product which results in multiple monomer and homodimer gel-shift complexes (* and †, respectively). Mixing studies with HA demonstrate that both Mix.1 and M11 form stable gel-shift complexes of intermediate mobility compared to Mix.1 monomers and homodimers (arrow).

METHODS. Embryos (stage 36) were stained for larval globin by immunohistochemistry with a monoclonal antibody as described²². Animal pole explants, with or without added activin (100 pM), were dissected from embryos injected with either BMP-4 or Mix.1 RNA (1 ng) at the one-cell stage, and analysed for *Xhox3* expression by using reverse transcription with polymerase chain reaction (RT-PCR) at stage 18. RNA extractions, first-strand cDNA synthesis and PCR analysis have been described⁴. Point mutations in Mix.1 cDNA were generated using the PCR overlap method²³. Proteins for gel-shift experiments were *in vitro* translated from synthetic RNA using rabbit reticulocyte lysate system (Promega).

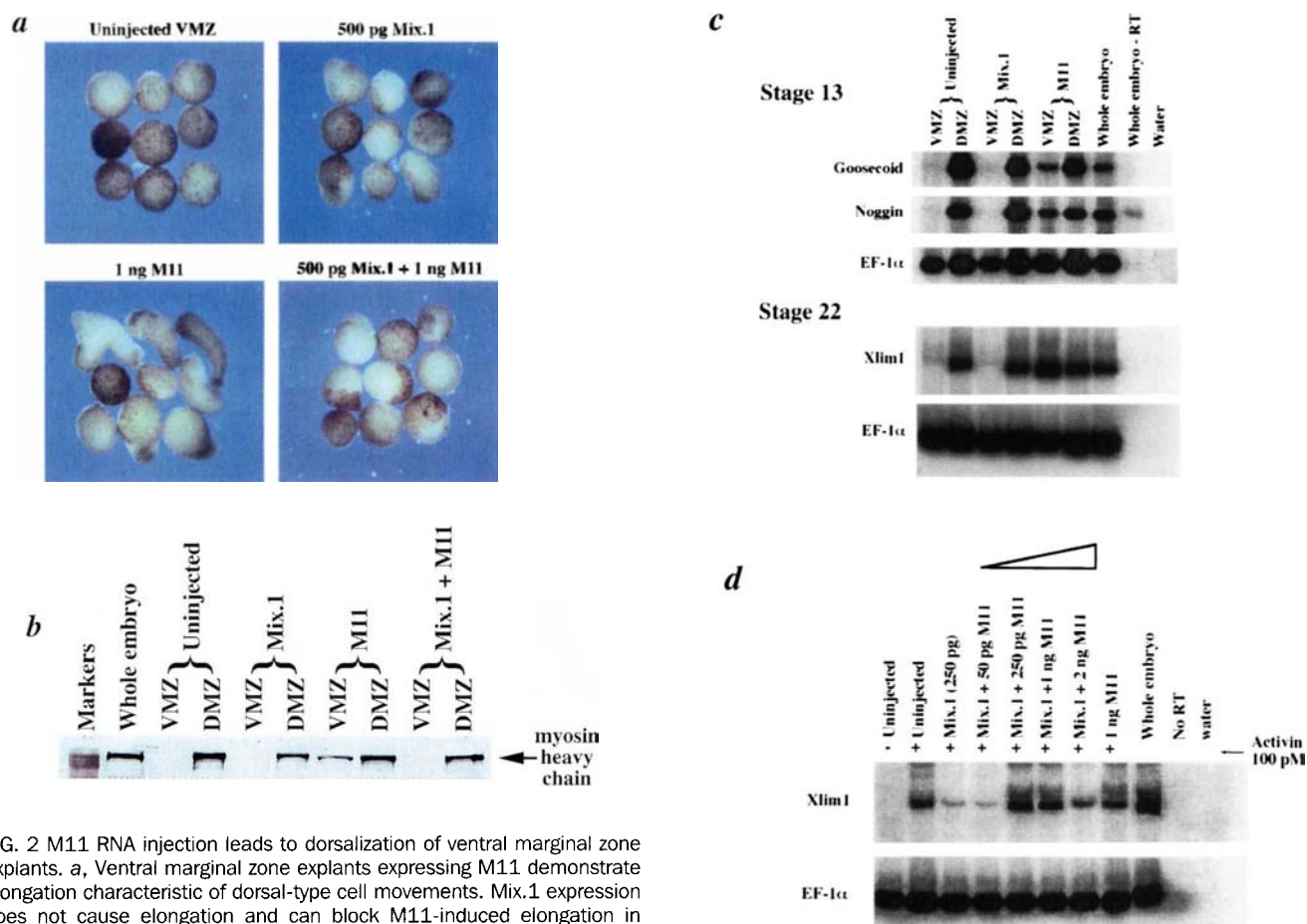


FIG. 2 M11 RNA injection leads to dorsalization of ventral marginal zone explants. **a**, Ventral marginal zone explants expressing M11 demonstrate elongation characteristic of dorsal-type cell movements. Mix.1 expression does not cause elongation and can block M11-induced elongation in ventral explants. **b**, Western-blot analysis for sarcomere myosin heavy chain demonstrates that ventral marginal zones from M11-injected embryos express this muscle-specific protein at stage 34. The specificity of M11 as an inhibitory mutant of Mix.1 is demonstrated by the finding that VMZs dissected from embryos injected with both Mix.1 and M11 do not elongate and do not express myosin heavy chain. **c**, RT-PCR analysis of dorsal-specific genes in ventral marginal zone explants dissected from embryos injected with M11. Injection of M11 results in expression of the dorsal-specific genes goosecoid and noggin at stage 13 and Xlim1 at stage 22. **d**, M11 functions as a dominant inhibitory mutant. Mix.1 expression in activin-treated caps led to a decrease in Xlim1 expression. Co-injection with M11 led to an increase in Xlim1 expression similar to levels seen with activin-treated animal pole explants. M11 by itself did not alter Xlim1 expression induced by activin.

METHODS. Embryos were injected in both blastomeres at the two-cell

stage with Mix.1 (500 pg), M11 (1 ng), or combinations of each. Marginal zones (presumptive mesoderm) were dissected at early gastrula stages ($10^{1/4}$) and cultured until the indicated stage. Western-blot analysis was performed at stage 34 using an anti-sarcomere myosin heavy-chain monoclonal antibody, MF20 (ref. 24), at 1/7,500 dilution and detected by alkaline phosphatase reagents as described⁴. Duplicate gels stained with Coomassie blue indicated similar amounts of total protein in each lane (not shown). Animal pole cells dissected from embryos injected with either Mix.1 (250 pg) or M11 (1 ng), or combinations of Mix.1 (250 pg) and M11 (50 pg to 2 ng), were explanted at stage 8, treated with activin protein (100 pM final concentration) and cultured until control stage 22. Xlim1 expression was assayed by RT-PCR analysis.

subpools until an individual cDNA was obtained. The nucleotide sequence of this clone was 98.5% identical to Mix.1, a homeobox protein originally isolated as an immediate early response gene to activin induction in *Xenopus* animal pole explants¹. Injection of Mix.1 RNA produces embryos that lack head structures and express globin throughout the dorsal-ventral axis (Fig. 1a). Mix.1 injected into dorsal blastomeres results in severe ventralization, whereas injection of the same dose on the ventral side has little effect on development (data not shown). Dorsal marginal zones (DMZs) dissected from embryos injected with Mix.1 expressed globin (82%, $n = 22$), whereas uninjected DMZs never express globin⁴ (0%, $n = 16$) (Fig. 1b). This demonstrates that Mix.1 is capable of re-specifying dorsal mesoderm to a ventral fate.

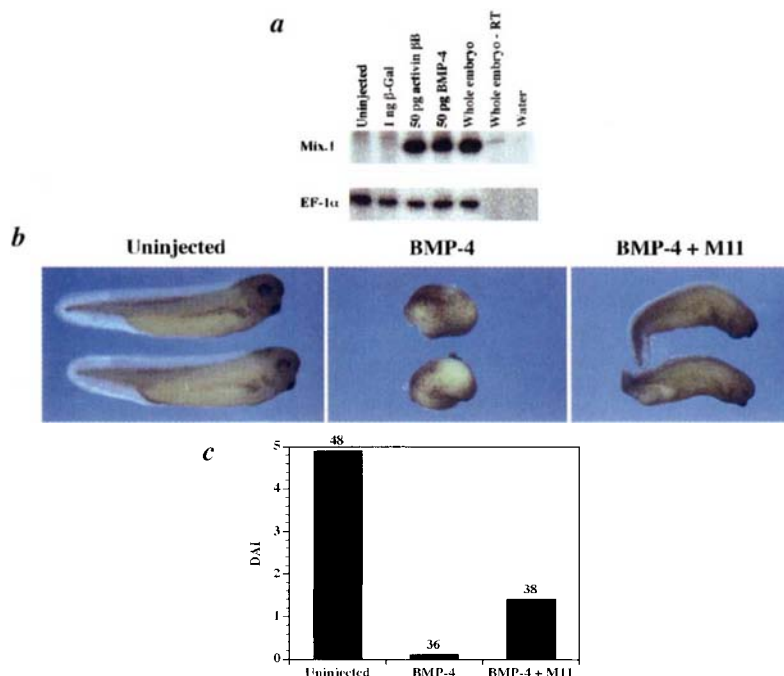
In animal pole explants, Mix.1 expression was not sufficient to induce Xhox3 (ref. 5) (Fig. 1c), Xbra (ref. 6), Xwnt8 (ref. 2), BMP-4 (refs 7,8) or noggin⁹ (data not shown). Although high doses of activin typically induce dorsal-type mesoderm¹⁰⁻¹², overexpression of Mix.1 in activin-treated caps led to expression of the ventral or

posterior markers, α -globin and Xhox3 (data not shown). These findings indicate that Mix.1 expression is not sufficient to induce mesoderm, but Mix.1 will pattern mesoderm to a ventral fate.

Mix.1 is related to the Pax family of homeodomain proteins, which interact with DNA as dimers^{13,14}. The mobility of Mix.1 monomers is distinguished from dimers with a half-site probe (Fig. 1d; oligo P₂). A deletion mutant of Mix.1, HA, encoding only the homeodomain, forms monomers and dimers similar to the full-length Mix.1. Mix.1-HA heterodimers have an intermediate gel-mobility-shift complex (Fig. 1d, arrow) compared to either HA or Mix.1 homodimers. We created a dominant inhibitory mutant (M11) of Mix.1 by substituting an isoleucine (at position 135) for a proline in the turn connecting α -helix II and α -helix III of the homeodomain. M11 binds to the P3 consensus site¹³ with ~100-fold-less affinity than Mix.1 (data not shown). Mix.1 and M11 can dimerize, as demonstrated by a stable gel-mobility-shift complex of M11 and HA (arrow); however, M11-HA heterodimers bind DNA with less affinity than Mix.1-HA heterodimers (Fig. 1d). These data are consistent with a model in which one wild-type

FIG. 3 BMP-4 induces the expression of Mix.1 in animal pole cells and the dominant inhibitory Mix.1 mutant (M11) blocks BMP-4 induced ventralization. **a**, BMP-4, like activin, can induce Mix.1 expression in animal pole cells. **b**, BMP-4 RNA injection leads to pronounced ventralization (dorso–anterior indices³ (DAI) of about zero). Co-injection of M11 RNA results in a restoration of a dorsal axis in these animals (DAI $\approx$ 2–4). **c**, The DAI scores for this experiment are summarized in the chart.

METHODS. For **a**, BMP-4 or activin β B (ref. 11) RNA (50 pg) was injected into the animal pole of embryos at the one-cell stage. Animal pole explants were dissected at stage 7–8 and cultured to stage 10^{1/2}. Mix.1 expression was determined by RT-PCR analysis. For **b**, embryos were injected at the two-cell stage into each blastomere with either BMP-4 (500 pg), M11 (1 ng), or BMP-4 and M11, and then scored after two days of development.



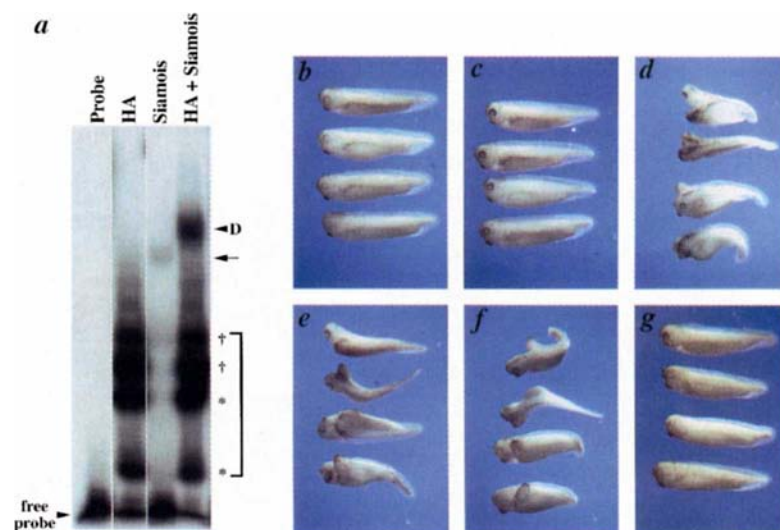
homeodomain is sufficient to interact with DNA and, once bound, the monomer can recruit other partners to the complex¹⁴.

The effect of the dominant inhibitory Mix.1 mutant was studied in marginal zone explants. In contrast to the typical round morphology, ventral marginal zone (VMZ) explants from M11-injected embryos underwent elongation characteristic of DMZ explants (Fig. 2a) and had a decrease in blood formation (data not shown). VMZs loaded with M11 also expressed the dorsal-specific mesodermal markers goosecoid¹⁵ and noggin⁹ (at stage 13; ref. 16), Xlim 1 (ref. 17) (stage 22; Fig. 2c) and myosin heavy chain (stage 34; Fig. 2b), whereas control VMZs did not. Mix.1 can prevent M11-induced elongation of VMZs (Fig. 2a) and sarcomere myosin heavy-chain expression (Fig. 2b). Co-injection of M11 RNA restored almost complete dorsal–ventral axes to Mix.1-loaded embryos, although defects in gastrulation were occasionally present (data not shown). Mix.1 expression in

DMZ explants led to expression of globin (Fig. 1b) but did not reduce the expression of goosecoid and noggin in DMZ explants (Fig. 2c). This implies that Mix.1 overexpression is not likely to perturb the early events of mesoderm induction in the dorsal marginal zone. The effect of M11 on Mix.1 activity was also analysed in animal pole explants. Activin treatment of animal caps induced expression of Xlim1 RNA, whereas activin-treated animal caps loaded with Mix.1 RNA had significantly decreased expression of Xlim1 (Fig. 2d). A 1:1 molar ratio of injected M11 and Mix.1 RNA restored control levels of Xlim1 expression in activin-treated caps. In transient transfection studies in NIH3T3 cells, Mix.1 was found to be a potent transcriptional activator using a P3-site reporter construct. At high doses, M11 blocked Mix.1-directed transactivation (data not shown). M11 thus has dorsalizing characteristics, distinctly opposite to Mix.1, and can function as a dominant inhibitory mutant.

FIG. 4 Interactions between Mix.1 and siamois. **a**, Siamois forms stable heterodimers with Mix.1. Gel-mobility-shift assay was performed with *in vitro* translated siamois and Mix.1 homeodomain, HA, using a tandem repeat P3 consensus site as a probe (see Fig. 2 for details). Siamois alone binds weakly to this probe (arrow) but forms stable heterodimers (D) with HA. **b–g**, Mix.1 expression rescues axis duplication induced by ectopic *siamois* expression. **b**, Control uninjected embryos. **c**, Mix.1 (50 pg) injection into ventral blastomere does not disrupt normal development. **d**, Axis duplication induced in sibling embryos by injection of 50 pg *siamois* RNA (50% complete secondary axis; $n = 36$). **e–g**, Co-injection of siamois (50 pg) and Mix.1 (e, 50 pg; f, 100 pg; g, 250 pg) RNA results in suppression of the secondary axis. Mix.1 (250 pg) virtually eliminates the second axis (0% complete secondary axis, 5% partial secondary axis; $n = 40$), although problems with gastrulation were sometimes noted.

METHODS. Full-length *siamois* RNA was *in vitro* translated using reticulocyte lysate (Promega). Gel-mobility-shift assay was performed with a ³²P-labelled P3 oligonucleotide probe¹³. Embryos were injected in one ventral–vegetal blastomere at the 8–16-cell stage and cultured for 2 d. Embryos were scored for axis duplication; complete axis duplication indicates the induction of eyes and cement gland in the secondary axis¹⁹.



Mix.1 and BMP-4 are each expressed in the ventral region of the embryo during gastrulation^{1,18} and both produce ventral phenotypes when injected into whole embryos. Injection of BMP-4 induces Mix.1 expression in animal pole cells (Fig. 3a). To determine whether BMP-4-induced ventralization requires Mix.1 function, we studied the effect of the dominant inhibitory mutant M11 on BMP-4 activity. Co-injection of 500 pg BMP-4 and 1 ng M11 RNA led to a partial rescue of the ventralized BMP-4 phenotype (Fig. 3b, c). These results indicate that Mix.1 is positioned downstream of BMP-4 in a signalling cascade that participates in ventral patterning.

We considered that Mix.1 might interact with other Pax-like homeodomain proteins during early embryogenesis. The homeodomain protein siamois is expressed on the dorsal side of the embryo and produces a duplicated axis when injected into ventral vegetal blastomeres of cleavage-stage embryos (Fig. 4b–d)¹⁹. Mix.1 and siamois interact *in vitro* to form stable heterodimers (Fig. 4a). Co-injection of increasing doses of Mix.1 RNA with 50 pg *siamois* RNA results in a dose-dependent reduction of the duplicated axis, such that a 5-fold excess of Mix.1 RNA will suppress secondary axis formation (Fig. 4e–g). This rescue, taken together with the structural similarity and the demonstration of heterodimer formation between Mix.1 and siamois, suggests a model in which these two proteins antagonize each other's function *in vivo*.

The Pax class of homeodomain proteins may participate in axial patterning by regulating gene expression through the action of homo- or heterodimers. A model of axial patterning invoking dimerization of homeodomain proteins is contingent on the co-expression of these genes *in vivo*. Based on *in situ* hybridization patterns, Mix.1 and siamois expression overlap in dorsal vegetal cells of the late blastula and early gastrula^{1,19}. Mix.1 homodimers lead to ventralization, and siamois may dorsalize by overriding the function of these homodimers. Goosecoid has also been shown to heterodimerize with Mix.1 (ref. 13), and may function similarly to siamois. Given that Mix.1 is a potent ventralizing factor and that Mix.1 can heterodimerize with other homeodomain proteins that have apparently competitive activities, we propose that Mix.1 plays a key role in establishing the dorsal–ventral axis. □

Received 4 March; accepted 20 May 1996.

- Rosa, F. M. *Cell* **57**, 965–974 (1989).
- Smith, W. C. & Harland, R. M. *Cell* **67**, 753–765 (1991).
- Kao, K. R. & Elinson, R. P. *Dev. Biol.* **127**, 64–77 (1988).
- Kelley, C., Yee, K., Harland, R. & Zon, L. *Dev. Biol.* **165**, 193–205 (1994).
- Ruiz-Altaba, A. & Melton, D. A. *Cell* **57**, 317–326 (1989).
- Smith, J. C., Price, B. M., Green, J. B., Weigel, D. & Herrmann, B. G. *Cell* **67**, 79–87 (1991).
- Dale, L., Howes, G., Price, B. M. J. & Smith, J. C. *Development* **115**, 573–585 (1992).
- Jones, C. M., Lyons, K. M., Lapan, P. M., Wright, C. V. E. & Hogan, B. L. M. *Development* **115**, 639–647 (1992).
- Smith, W. C. & Harland, R. M. *Cell* **70**, 829–840 (1992).
- Green, J. B. & Smith, J. C. *Nature* **347**, 391–394 (1990).
- Thomsen, G. et al. *Cell* **63**, 485–493 (1990).
- Green, J. B., New, H. V. & Smith, J. C. *Cell* **71**, 731–739 (1992).
- Wilson, D., Sheng, G., Lecuit, T., Dostatni, N. & Desplan, C. *Genes Dev.* **7**, 2120–2134 (1993).
- Wilson, D. S., Guenther, B., Desplan, C. & Kuriyan, J. *Cell* **82**, 709–719 (1995).
- Niehrs, C., Steinbeisser, H. & De Robertis, E. M. *Science* **263**, 817–820 (1994).
- Nieuwkoop, P. D. & Faber, J. *Normal Table of Xenopus laevis* (Daudin, North Holland, Amsterdam, 1967).
- Taira, M., Jamrich, M., Good, P. J. & Dawid, I. B. *Genes Dev.* **6**, 356–366 (1992).
- Fainsod, A., Steinbeisser, H. & De Robertis, E. M. *EMBO J.* **13**, 5015–5025 (1994).
- Lemaire, P., Garrett, N. & Gurdon, J. B. *Cell* **81**, 85–94 (1995).
- Strauss, E. C., Andrews, N. C., Higgs, D. R. & Orkin, S. H. *Molec. cell. Biol.* **12**, 2135–2142 (1992).
- Le Roux, I., Joliot, A. H., Block-Gallego, E., Prochiantz, A. & Volovitch, M. *Proc. natn. Acad. Sci. U.S.A.* **90**, 9120–9124 (1993).
- Hemmati-Brivanlou, A. & Harland, R. M. *Development* **106**, 611–617 (1989).
- Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. & Pease, L. R. *Gene* **77**, 51–59 (1989).
- Bader, D., Masaki, T. & Fischman, D. A. *J. Cell Biol.* **95**, 763–770 (1982).

ACKNOWLEDGEMENTS. We thank the members of our laboratory for discussions and suggestions; S. Orkin, B. Paw, A. Perkins, D. Ransom, R. Shivdasani, J. Turpen and M. Weiss for comments on the manuscript; P. Smith and R. Harland for the vector and for insight into expression cloning; C. H. Katagiri for globin monoclonal antibody; P. Lemaire and G. Thomsen for the *siamois* and BMP-4 cDNA constructs, respectively; and D. Melton and A. Hemmati-Brivanlou for advice on *Xenopus* biology. This work was supported by a grant from the NIH (L.I.Z.). L.I.Z. is an assistant investigator of HHMI; P.E.M. is an Associate of HHMI.

CORRESPONDENCE and requests for materials should be addressed to L.I.Z. (e-mail: zon@rascal.med.harvard.edu).

The mouse *Engrailed-1* gene and ventral limb patterning

Cynthia A. Loomis^{*†}, Esther Harris[†],
Jacques Michaud[‡], Wolfgang Wurst[§],
Mark Hanks^{†¶} & Alexandra L. Joyner^{†¶}

^{*} Ronald O. Perleman Department of Dermatology, and [¶] Departments of Cell Biology, and Physiology and Neuroscience, New York University Medical School, 550 First Avenue, New York, New York 10016, USA

[†] Developmental Genetics Program, Skirball Institute of Biomolecular Medicine, New York University Medical School, 540 First Avenue, New York, New York 10016, USA

[‡] Institut D'Embryologie Cellulaire et Moléculaire Du Collège De France et du CNRS, 49 bis, avenue de la Belle Gabrielle, 94736 Nogent-sur-Marne, France

[§] GSF-Research Center, Ingolstaedter Landstrasse 1, Oberrheissheim-Munich, D-85758, Germany

DURING vertebrate limb development, positional information must be specified along three distinct axes. Although much progress has been made in our understanding of the molecular interactions involved in anterior–posterior and proximal–distal limb patterning, less is known about dorsal–ventral patterning^{1–3}. The genes *Wnt-7a* and *Lmx-1*, which are expressed in dorsal limb ectoderm and mesoderm, respectively, are thought to be important regulators of dorsal limb differentiation^{4–6}. Whether a complementary set of molecules controls ventral limb development has not been clear. Here we report that *Engrailed-1*, a homeodomain-containing transcription factor expressed in embryonic ventral limb ectoderm^{7,8}, is essential for ventral limb patterning. Loss of *Engrailed-1* function in mice results in dorsal transformations of ventral paw structures, and in subtle alterations along the proximal–distal limb axis. *Engrailed-1* seems to act in part by repressing dorsal differentiation induced by *Wnt-7a*, and is essential for proper formation of the apical ectodermal ridge.

Expression of *Engrailed-1* (*En-1*) in the ventral ectoderm of developing limbs suggests a role in dorsal–ventral (D–V) limb patterning, as tissue transplantation studies show that the ectoderm plays an important part in this process^{9–12}. We have analysed the distribution of epidermal appendages in mice that have a null mutation in *En-1*, *En-1^{hd}*, because these structures provide unambiguous markers for dorsal and ventral limb differentiation. In contrast to the dorsally positioned nails of newborn wild-type mice, the nails on digits 1–4 of *En-1^{hd}* mice were circumferential ($n = 44$ limbs), apparently supplanting the distal, ventral pads (Fig. 1a–e). In addition, serial sections of four *En-1* mutant limbs showed that dorsally restricted hair follicles were present along ventral as well as dorsal digit surfaces of *En-1^{hd}* mutant mice, whereas ventrally restricted eccrine glands were largely absent (Fig. 1d, e). Tendons and bones, two purely mesodermally derived structures, were also affected. Unlike wild-type digits, *En-1^{hd}* digits flexed dorsally (38 of 44 limbs) rather than ventrally, suggesting that there may be ventral tendon abnormalities (Fig. 1d, e). Histologically, the ventral tendons of the digit flexor muscles in two mutant limbs analysed were absent or poorly developed (Fig. 1f, g). Ventral skeletal elements, such as the calciform and sesamoid cartilages, were lost or incompletely formed in *En-1^{hd}* mutant paws ($n = 13$ limbs) (Fig. 1h, i). Thus *En-1^{hd}* mice displayed aspects of double dorsal paws.

Phenotypic studies initially focused on newborn limbs because *En-1^{hd}* mutants usually die shortly after birth as a result of a severe deletion of mid-hindbrain structures¹³. Subsequent analysis of adult mice homozygous for two other *En-1* mutant alleles, *En-1^{dki}* and *En-1^{2ki}*, and of rare *En-1^{hd}* mice surviving to three weeks

[†] Present address: The Proctor and Gamble Company, Miami Valley Laboratories, Route 27, 11810 East Miami River Road, Ross, Ohio 45061, USA.

of age, confirmed the functional importance of *En-1* in specifying ventral limb structures. Many *En-1^{dki/dki}* and all *En-1^{2ki/2ki}* mice survive to adulthood owing to expression of *Drosophila engrailed* or mouse *En-2*, respectively, from the *En-1* locus and rescue of the *En-1^{hd/hd}* brain anomalies¹⁴ (M.H. *et al.*, manuscript in prepara-

tion). The D-V limb phenotypes of *En-1^{hd/hd}* and *En-1^{dki/dki}* mutants were similar, showing circumferential nails ($n = 32$), ventral hairs ($n = 32$), and loss of eccrine glands and sesamoid bones ($n = 4$) (Fig. 2a-c, e, f). The *En-1^{2ki/2ki}* mice displayed a less severe phenotype (Fig. 2d, h). The limbs appeared normal at birth,

FIG. 1 Loss of *En-1* function results in the dorsal transformation of ventral structures and ventral duplication of distal structures in *En-1^{hd/hd}* newborn limbs. Ventral views of wild type (a) and *En-1^{hd/hd}* (b, c) forepaws illustrating the circumferential mutant nails (white arrowhead) and occasional ectopic ventral digits (6, 7) that emanate from the base of proximal mutant pads. d, Haematoxylin and eosin stained sagittal section of a distal wild-type paw, dorsal side up, showing the dorsally restricted, pale-staining nail plate (n) and ventrally restricted paw pads (black brackets) which contain immature eccrine glands. e, A similar *En-1^{hd/hd}* section demonstrating the presence of nail plates on both ventral and dorsal digit surfaces (n), absence of distal-most paw pads (asterisk) and loss of eccrine glands in proximal pads. Serial sections of *En-1^{hd/hd}* digits confirmed that dermal papillae (see inset) were present at the base of epithelial structures in the ventral as well as dorsal dermis (open arrowheads), identifying these structures as hair follicles not eccrine glands. Note the ventral (palmar) flexion of the wild-type paw in contrast to the dorsal flexion of the *En-1^{hd/hd}* paw. f, g, Cross-sections illustrating the ventral flexor digitorum profundus (central) tendon and flexor digitorum superficialis (lateral) tendons in the ventral mesenchyme of wild-type digits (f) and the poorly developed flexor digitorum profundus tendon, which resembles the dorsal extensor digitorum longus tendon, and absent flexor digitorum superficialis tendons in *En-1^{hd/hd}* digits (g) (white brackets demarcate the ventral tendons). h, i, Skeletal preparations demonstrating the lack of metatarsal-associated sesamoid cartilage (black arrow) in an *En-1^{hd/hd}* digit (i) compared to an analogous wild-type digit (h). Also note that the asymmetric ventral protrusions at the proximal end of wild-type phalanges are absent from mutant phalanges.

METHODS. Heterozygous mice carrying the *En-1^{hd}* mutation were intercrossed and newborn pups genotyped by Southern blot analysis using tail sample DNA¹³. For histological analysis, limbs were fixed in formalin, embedded in paraffin, and sections (5 μ m) cut and stained with haematoxylin and eosin. Alizarin red/alcan blue skeletal staining was performed as described²¹. Scale bar, 250 μ m.

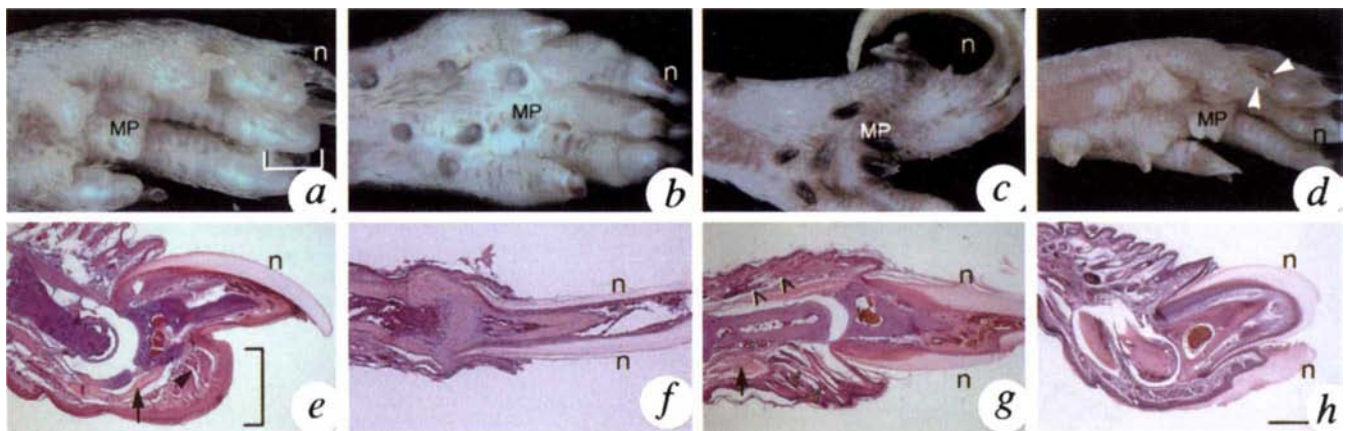
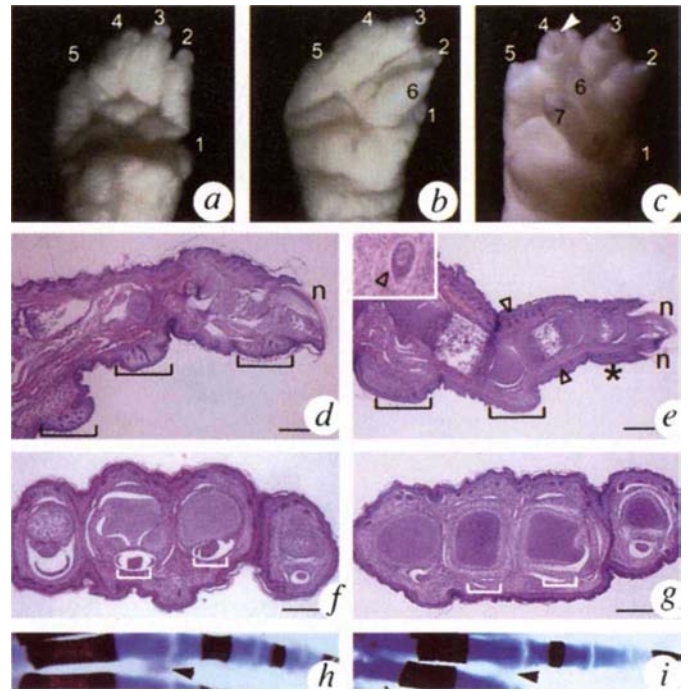


FIG. 2 Adult mice homozygous for three distinct *En-1* mutant alleles show dorsalization of the ventral paws. a-d, Ventral views of adult hindpaws, and e-h, haematoxylin and eosin stained sagittal sections of distal digits with dorsal surfaces up and distal tips to the right. a, e, Limbs of a wild-type mouse 4-6 months old illustrating mature dorsal nails (n), distal phalangeal pads (brackets), and proximal metatarsal pads (MP). Mature convoluted eccrine tubules are present within pad mesenchyme (e, black arrowheads). The tendon of the flexor digitorum profundus muscle can be seen inserting onto the proximal portion of the distal-most phalange (e, black arrow). b, f, Distal limbs of a 3-week-old *En-1^{hd/hd}* mouse, demonstrating the presence of ectopic ventral hairs. Note the circumferential nails and loss of distal pads and associated eccrine glands. Proximal metatarsal pads are present (MP), but possess few if any eccrine glands, and the distal side appears to differentiate into nail, histologically (data not shown). c, g, Distal limbs of a 14-month-old *En-1^{dki/dki}* mouse, illustrating cylindrical nails (n), normal dorsal and ectopic ventral hairs (open arrowheads), and loss of

distal pads and eccrine glands. Fusion of the digits, as shown here, is variable. The pigmented metatarsal pads (MP) are elongated and hardened, resembling nails (c and data not shown). The flexor digitorum profundus tendons are present (g, black arrow), but appear to end blindly in the distal soft tissue of most *En-1^{dki/dki}* digits. d, h, Distal limbs of a 3-month-old *En-1^{2ki/2ki}* mouse, illustrating the closely juxtaposed ventral and dorsal nails (d, white arrowheads). The unusual pad/nail hybrid structures associated with the metatarsal pads (MP) are clearly visible.

METHODS. Whole mounts and histological sections were prepared as described in Fig. 1, except that limbs were decalcified with Decalcifier II (Surgi-Path) to allow sectioning through bones. The *En-1^{dki}* and *En-1^{2ki}* alleles were generated by gene targeting to bring the *Drosophila en* and mouse *En-2* protein-coding sequences, respectively, under the control of the *En-1* regulatory sequences, and to inactivate *En-1* protein function¹⁴ (M.H. *et al.*, manuscript in preparation). Scale bars, 200 μ m.

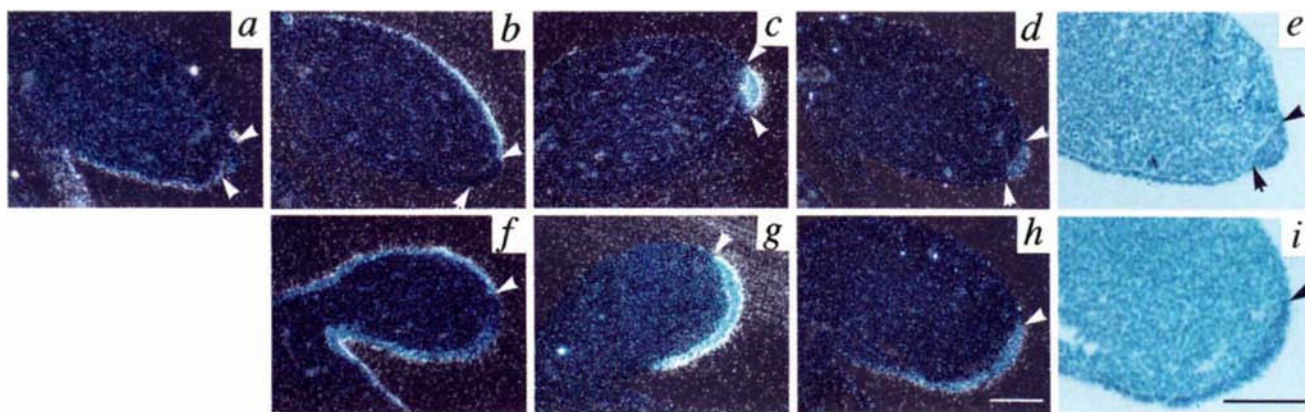


FIG. 3 Loss of *En-1* function in the limbs results in altered expression of dorsal- and AER-specific ectoderm markers. Dark-field photographs of RNA *in situ* hybridizations of limb sagittal sections from wild type (a–e) and *En-1*^{hd/hd} (f–i) 10.5 d.p.c. embryos. Dorsal up, distal to the right; white arrowheads mark the morphological borders of the AER. a, *En-1* antisense probe illustrating the restricted ventral ectoderm expression. b, f *Wnt-7a* probe demonstrating the restricted dorsal ectoderm expression in wild-type embryos, but expanded ventral expression in *En-1*^{hd/hd} mutants. Note *Wnt-7a* expression is excluded from the wild-type AER and from a region at the distal tip of the mutant limb. c, g, *Fgf-8*; d, h, *Bmp-2*, probes showing their striking delineation of the AER in wild-type limbs, but marked ventral

expansion in *En-1*^{hd/hd} mutants. e, i, Bright-field photographs of a higher power magnification of the toluidine blue-stained sections shown in d, h, showing the thickened, sharply demarcated wild-type AER at the distal tip of the limb (arrowheads) and the thinner *En-1*^{hd/hd} mutant AER, which retains a sharply demarcated dorsal border (arrowhead). The ventral portion of the AER, however, appears extended ventrally with gradual thinning of the ectoderm more proximally. METHODS. Embryos (10.5 d.p.c.) from heterozygous *En-1* matings were genotyped using DNA from yolk-sac samples and fixed in 4% paraformaldehyde for RNA *in situ* analysis²². ³⁵S-antisense probes specific for *En-1*, *Wnt-7a*, *Fgf-8* and *Bmp-2* were made as described^{13,4,16,23}. Scale bars, 100 μ m.

but after a few weeks ventral nails formed opposite the dorsal nails ($n = 64$ limbs).

In addition to these D–V patterning defects, the *En-1* mutants also demonstrated proximal–distal limb abnormalities. Extra digits emanated from the ventral–proximal paws of 10–15% of *En-1*^{hd/hd} mutants¹³ ($n = 116$ limbs) (Fig. 1a–c). By 6 weeks of age the more proximal pads of *En-1*^{dkid/dki} and *En-1*^{2ki/2ki} mutant mice had developed unusual nail-pad hybrid structures, reminiscent of digit tips ($n = 96$). Thus structures typically limited to the distal

limb (nails and digits) also developed proximoventrally in *En-1* mutants.

To investigate the molecular mechanisms by which *En-1* regulates ventral limb patterning, we analysed morphology and gene expression in limb buds of 10.5 day post coitum (d.p.c.) wild-type and *En-1*^{hd/hd} embryos (Fig. 3). At this stage, both forelimb and hindlimb buds are visible, and the AER is well developed on the forelimbs. Limb sections stained with toluidine blue showed that the *En-1*^{hd/hd} AER was thinner than in the wild type, and extended ventrally and proximally (Fig. 3h, i). Consistent with these morphological changes, the expression of two AER marker genes, *Fgf-8* and *Bmp-2* (refs 15,16) demonstrated a ventral expansion in mutant limb buds (Fig. 3c, d, g, h, and data not shown).

Wnt-7a expression was also examined because this gene is required for normal dorsal limb mesenchyme differentiation⁴. In contrast to the dorsally restricted expression of *Wnt-7a* in wild-type limb buds¹⁷, *Wnt-7a* in *En-1*^{hd/hd} limbs was expressed in ventral as well as dorsal limb ectoderm (Fig. 3b, f). *Wnt-7a* expression, however, was not upregulated in the distal-most ectoderm of the mutant limbs.

These studies tell us more about the genetic pathways regulating D–V patterning of the vertebrate limb. A model for limb development has been discussed in which ventral-type differentiation is the default pathway and dorsal determining molecules are required to override ubiquitously expressed ventral specifying genes^{2,4–6}. This model, however, does not readily account for both the double-ventral phenotype of *Wnt-7a* mutant mice⁴ and the double-dorsal phenotype observed in *En-1* mutants. A more likely model is that distinct and competing molecular mechanisms lead to normal D–V patterning in the limb². In the ventral limb, *En-1* appears to promote ventral-type skin differentiation (eccrine glands) and inhibit dorsal-type differentiation (hairs and nails), probably in part through its suppression of the dorsalizing gene, *Wnt-7a*.

Our studies also provide new information on AER formation. The AER, like the non-AER

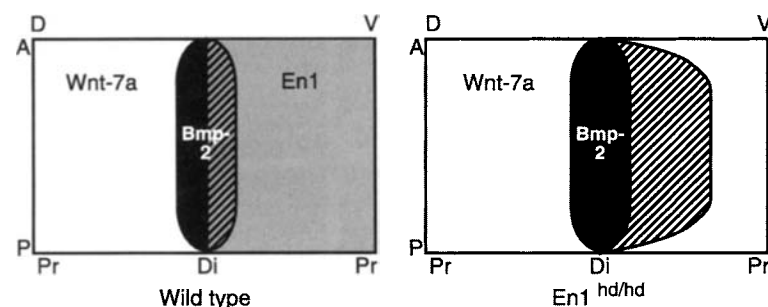


FIG. 4 Ectoderm domains of a, wild-type and b, *En-1* mutant embryonic mouse limbs. The early limb ectoderm, which is opened up and flattened so that dorsal (D) is on the left and ventral (V) is on the right, can be divided into four distinct domains based on gene expression in wild-type embryos. a, Limb expression of *En-1*, *Wnt-7a* and *Fgf-8*/*Bmp-2* in normal embryos. The dorsal non-AER ectoderm expresses only *Wnt-7a* (white). The ventral non-AER ectoderm expresses only *En-1* (grey). The dorsal AER ectoderm expresses *Fgf-8*/*Bmp-2* (black), whereas the ventral AER ectoderm coexpresses *Fgf-8*/*Bmp-2* with *En-1* (black and grey hatching). In mice lacking a functional *En-1* gene (b), the ventral as well as dorsal non-AER ectoderm express *Wnt-7a* (white). The *Fgf-8*/*Bmp-2* expression domain (solid black and black hatching) is expanded ventrally and proximally, suggesting a ventral expansion of the AER, consistent with the morphological data. The expanded but thinned AER in the *En-1*^{hd/hd} limbs also appears to be subdivided into two domains based on *Wnt-7a* expression: a dorsal *Wnt-7a*-negative region (solid black), and a ventral *Wnt-7a*-positive region (black and white hatching). These regions may correspond to the dorsal and ventral AER domains normally distinguished by *En-1* expression, suggesting that *Wnt-7a* in *En-1*^{hd/hd} mice, like *En-1* in wild types, is expressed in the ventral AER as well as non-AER ventral ectoderm. Alternatively, the *Wnt-7a*-negative (*Fgf-2*/*Bmp-2*-positive) domain may reflect the entire AER domain, indicating *En-1* is required to repress *Fgf-8* and *Bmp-2* in non-AER ventral ectoderm.

ectoderm, can be divided into two distinct domains (Fig. 4). The dorsal AER is characterized by the presence of AER-specific markers, whereas the ventral AER is distinguished by the additional expression of *En-1*. Loss of *En-1* function appears to allow a ventral expansion of the AER. In this situation, the *Wnt-7a*-negative region at the distal tip of the mutant limbs might be analogous to the dorsal wild-type AER, whereas the ectoderm expressing *Wnt-7a* in addition to AER-specific markers might represent the ventral AER. Alternatively, the distal *Wnt-7a*-negative ectoderm might demarcate the entire functional domain of the AER, but several pieces of evidence suggest that this is not the case. Both morphological criteria and gene-expression

data suggest that the AER extends beyond the ventral boundary of the *Wnt-7a*-negative domain. Furthermore, preliminary studies suggest a parallel proximodistal expansion of progress zone markers (data not shown). Finally, limb structures that normally develop only distally were duplicated proximodistally in *En-1* mutant mice, a phenotype consistent with a functional expansion of the AER. Thus our data indicate that *En-1* is required for delineating the ventral AER boundary and for restricting expression of signalling molecules, such as *Fgf-8* and *Bmp-2*, to the distal-most ectoderm, a function reminiscent of *engrailed's* role in compartment border formation in *Drosophila*^{18–20}. □

Received 20 March; accepted 20 May 1996.

1. Johnson, R. L., Riddle, R. D. & Tabin, C. *Curr. Opin. Genet. Dev.* **4**, 535–542 (1994).
2. Tickle, C. *Curr. Opin. Genet. Dev.* **5**, 478–484 (1995).
3. Martin, G. R. *Nature* **374**, 410–411 (1995).
4. Parr, B. A. & McMahon, A. P. *Nature* **374**, 350–353 (1995).
5. Riddle, R. D. et al. *Cell* **83**, 631–640 (1995).
6. Vogel, A., Rodriguez, C., Warnken, W. & Izpisua Belmonte, J. C. *Nature* **378**, 716–720 (1995).
7. Davis, C. A. & Joyner, A. L. *Genes Dev.* **2**, 1736–1744 (1988).
8. Joyner, A. L. & Martin, G. R. *Genes Dev.* **1**, 29–38 (1987).
9. MacCabe, J. A., Erick, J. & Saunders, J. W. Jr *Dev. Biol.* **39**, 69–82 (1974).
10. Patau, M.-P. in *Vertebrate Limb and Somite Morphogenesis* (eds Ede, D. A., Hinchliffe, J. R. & Balls, M.) 257–266 (Cambridge Univ. Press, 1977).
11. Geduspan, J. S. & MacCabe, J. A. *Dev. Biol.* **124**, 398–408 (1987).
12. Geduspan, J. S. & MacCabe, J. A. *Anat. Rec.* **224**, 79–87 (1989).
13. Wurst, W., Auerbach, A. B. & Joyner, A. L. *Development* **120**, 2065–2075 (1994).
14. Hanks, M., Wurst, W., Anson-Cartwright, L., Auerbach, A. B. & Joyner, A. L. *Science* **269**, 679–682 (1995).

15. Lyons, K. M., Pelton, R. W. & Hogan, B. L. M. *Development* **109**, 833–844 (1990).
16. Crossley, P. H. & Martin, G. R. *Development* **121**, 439–451 (1995).
17. Parr, B. A., Shea, M. J., Vassileva, G. & McMahon, A. P. *Development* **119**, 247–261 (1993).
18. Sanicola, M., Sekelsky, J., Elson, S. & Gelbart, W. M. *Genetics* **139**, 745–756 (1995).
19. Zecca, M., Basler, K. & Struhl, G. *Development* **121**, 2265–2278 (1995).
20. Tabata, T., Schwartz, C., Gustavson, E., Ali, Z. & Kornberg, T. B. *Development* **121**, 3359–3369 (1995).
21. Lufkin, T. et al. *Nature* **359**, 835–841 (1992).
22. Wilkinson, D. G., Bailes, J. A. & McMahon, A. P. *Cell* **50**, 79–88 (1987).
23. Lyons, K. M., Pelton, R. W. & Hogan, B. L. M. *Genes Dev.* **3**, 1657–1668 (1989).

ACKNOWLEDGEMENTS. We thank S. Chu for histology and advice; C.-X. Tong for technical help; K. Briegel for one *En-1^{tm1}* mouse that survived three weeks; G. Martin, A. McMahon and B. Hogan for sharing plasmids; and L. Niswander, G. Fischell and W. M. O'Guin for advice and comments on the manuscript.

CORRESPONDENCE and requests for materials should be addressed to A.L.J. (e-mail: joyner@saturn.med.nyu.edu).

Influence of dendritic structure on firing pattern in model neocortical neurons

Zachary F. Mainen* & Terrence J. Sejnowski

Howard Hughes Medical Institute, Computational Neurobiology Laboratory, Salk Institute for Biological Studies, La Jolla, California 92037, and Department of Biology, University of California, San Diego, La Jolla, California 92093, USA

NEOCORTICAL neurons display a wide range of dendritic morphologies, ranging from compact arborizations to highly elaborate branching patterns¹. *In vitro* electrical recordings from these neurons have revealed a correspondingly diverse range of intrinsic firing patterns, including non-adapting, adapting and bursting types^{2,3}. This heterogeneity of electrical responsiveness has generally been attributed to variability in the types and densities of ionic channels. We show here, using compartmental models of reconstructed cortical neurons, that an entire spectrum of firing patterns can be reproduced in a set of neurons that share a common distribution of ion channels and differ only in their dendritic geometry. The essential behaviour of the model depends on partial electrical coupling of fast active conductances localized to the soma and axon and slow active currents located throughout the dendrites, and can be reproduced in a two-compartment model. The results suggest a causal relationship for the observed correlations between dendritic structure and firing properties^{3–7} and emphasize the importance of active dendritic conductances in neuronal function^{8–10}.

We began with a compartmental model used in a previous study of spike initiation⁸. This model included a low density of Na⁺ channels in the soma and dendrites¹¹ and a high density in the axon hillock and initial segment^{12,13}. Fast K⁺ channels were present in the axon and soma but excluded from the dendrites. To extend the

model from single spikes to spike trains, slow K⁺ channels responsible for the spike afterhyperpolarization (AHP) and control of repetitive firing^{14,15} (both calcium-dependent and voltage-dependent) were added to the soma and dendrites, along with one type of high-threshold Ca²⁺ channel.

Dendritic arborizations were designed using reconstructions of neocortical neurons with a variety of morphologies (Fig. 1). When simulated with the same channel types and densities, a range of distinct firing patterns was produced. Firing patterns correlated strongly with the extent of arborization and, to the degree that it was correlated with morphology, with the cell layer. Smooth stellates ($n = 4$), with the smallest dendritic arborizations, produced spike trains with deep, monophasic AHPs and very weak spike-frequency adaptation (Fig. 1a). Layer 4 spiny stellates ($n = 5$), with somewhat larger dendritic trees, gave adapting spike trains (Fig. 1b). Still more extensive layer 2 and 3 pyramidal neurons (Fig. 1c) showed spike after-depolarizations (ADPs) along with doublet or burst firing. The largest cells, layer 5 pyramidal neurons ($n = 7$), produced repetitive bursting spike trains (Fig. 1d). Although the specific set of patterns produced across this set of cells could be altered by changing the channel densities, the correlation between dendritic structure and firing pattern was insensitive to variations in the choice of channel kinetics and channel densities, as long as the basic regional segregation of fast and slow channel types was maintained. The results also seemed robust to heterogeneity of channel densities within the dendritic tree (as suggested by physiological studies^{16,17}), although such scenarios were not examined exhaustively.

To facilitate analysis of the role of dendritic electrical structure in shaping firing behaviour, a reduced two-compartment version of the model¹⁸ was examined (Fig. 2a). In this reduced model, electrical structure was determined by just two parameters: the ratio of axo-somatic area to dendritic membrane area (ρ) and the coupling resistance between axo-somatic and dendritic compartments (κ). These parameters were systematically varied while holding channel densities constant. When uncoupled ($\kappa \rightarrow \infty$), simple stereotyped oscillatory patterns were generated in the two compartments (Fig. 2b), with the axo-somatic compartment having a much higher intrinsic frequency owing to its relatively

* Present address: Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA.

rapid channel kinetics. When the dendrite and axon–soma were fully coupled ($\kappa \rightarrow 0$), alteration of ρ affected the size of AHP and the amount of spike-frequency adaptation, but produced a limited set of firing properties (Fig. 2c).

A weak to moderate coupling strength was necessary to produce the full range of firing patterns, including spike bursts and ADPs. With such partial coupling, small changes in ρ or κ produced dramatic changes in firing pattern (Fig. 2d, e). Dendritic Na^+ channels were critical to the generation of bursts and ADPs^{6,19,20} (Fig. 3a–c). These channels promoted propagation of spikes from the axon–soma into the dendrite, by prolonging the depolarization following a spike²¹. After the soma and axon had repolarized, current returned from the dendrite to produce a late depolarizing transient (the ADP), with somatic spike bursts occurring when ADPs were above threshold²². Decreasing κ or ρ suppressed bursting and ADPs by reducing the delayed depolarization associated with the dendritic spike (Fig. 3d, e). In addition, factors that affected the amplitude and duration of the dendritic spike, including the degree of Na^+ -channel inactivation and the strength of dendritic K^+ currents, similarly affected bursting and ADPs. Thus, dendritic Ca^{2+} channels could both increase (directly through their depolarizing current) and decrease (indirectly by activating Ca^{2+} -dependent K^+ channels) these phenomena. However, the effects of electrical structure on firing pattern and voltage trajectory were not strictly dependent on Ca^{2+} , as similar results could be obtained with a channel set lacking I_{Ca} and I_{KCa} , in

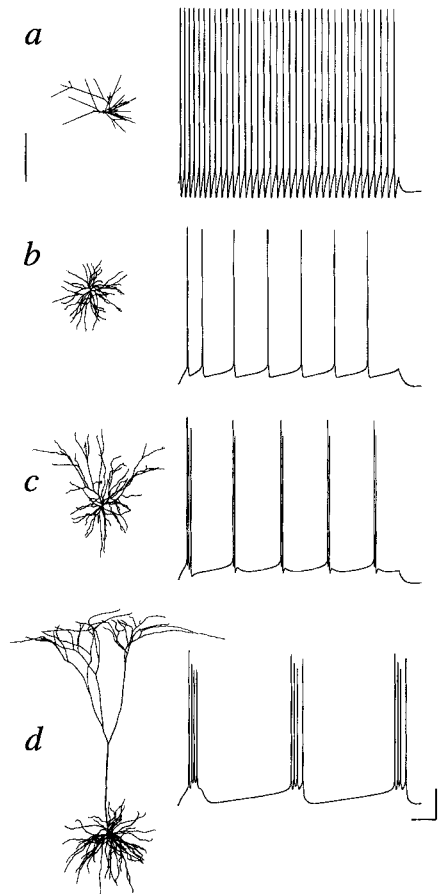
which interaction between I_{Na} and I_{KCa} gave rise to bursts (not shown).

To relate the reconstructed multicompartmental models back to the reduced model we examined their electrotonic structure in more detail (Fig. 4). The neurons examined varied widely in both their dendritic membrane area and the degree of electrical attenuation between soma and dendrites. This can be seen in histograms of the steady-state electrical impedances between soma and dendritic compartments (see Fig. 4 legend). This metric is analogous to the parameter κ of the two-compartment model. Fast non-adapting spiking occurred for neurons with the smallest dendritic area and transfer impedances; adapting spike trains accompanied moderate dendritic area and transfer impedance; and bursting was associated with the largest dendritic area and transfer impedances. These findings are consistent with comparisons of electrical structure of ‘thick’ (bursting) and ‘slender’ (non-bursting) layer 5 cells²³ and with similarities in the developmental time course of electrical structure and electrophysiological properties of neocortical neurons²⁴.

Our results demonstrate that the electrotonic structure of a neuron shapes the dynamic interactions between non-uniformly distributed ion channels, and may thereby control the pattern of repetitive firing and the interspike membrane-potential trajectory. Heterogeneity of dendritic structure can thereby explain several aspects of the heterogeneous firing properties of neocortical neurons parsimoniously in terms of their anatomical variety.

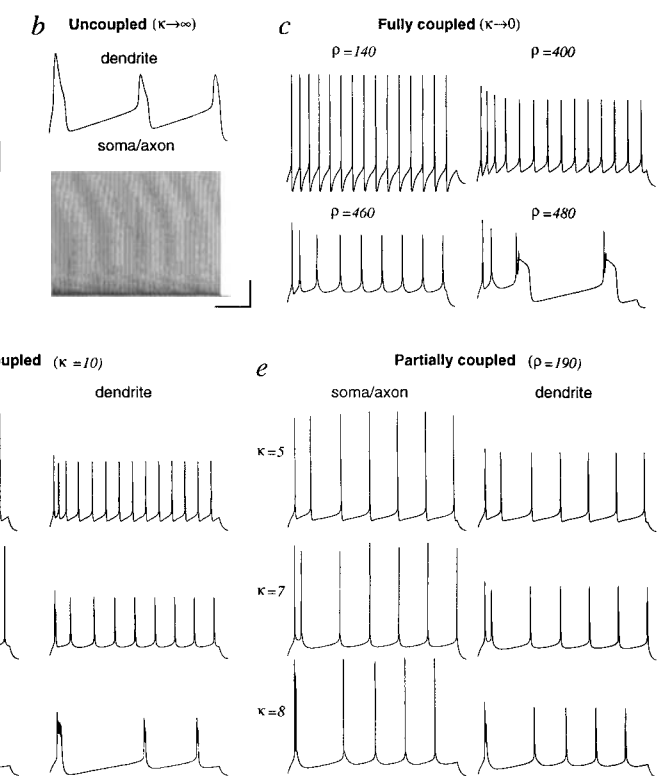
FIG. 1 Distinct firing patterns in model neurons with identical channel distributions but different dendritic morphology. Digital reconstructions of dendritic arborizations of neurons from rat somatosensory cortex (a) and cat visual cortex (b–d). a, Layer 3 aspiny stellate. b, Layer 4 spiny stellate. c, Layer 3 pyramid. d, Layer 5 pyramid. Somatic current injection (50, 70, 100, 200 pA for a–d, respectively) evoked characteristic firing patterns. a shows only the branch lengths and connectivity whereas b–d show a two-dimensional projection of the three-dimensional reconstruction. Scale bars: 250 μm (anatomy), 100 ms, 25 mV.

METHODS. Standard compartmental modelling techniques²⁵ were used to simulate spatially extended neurons with passive electrical structure, four voltage-dependent currents: fast Na^+ , I_{Na} (refs 8, 27); fast K^+ , I_{K} (refs 8, 27); slow non-inactivating K^+ , I_{Ks} (ref. 28); and high-voltage activated Ca^{2+} , I_{Ca} (ref. 29) and one Ca^{2+} -dependent current, I_{KCa} (ref. 30). All dendritic branches were divided into cylindrical compartments with a maximum length of 50 μm . The dendritic membrane area of spiny neurons was increased to account for spines (adding 0.83 μm^2 per linear μm of dendrite). An axon, which was not present in the reconstructed anatomy, was attached to the soma of each cell⁸. The axon consisted of a conical hillock (10 μm long), tapering to one-quarter width to a cylindrical initial-segment region (15 μm) followed by 5 myelinated internodes (100 μm) separated by nodal segments. We took into account an observed correlation between soma diameter and initial segment diameter³⁰, by scaling the initial segment diameter as a function of the soma area. All currents were calculated using conventional Hodgkin–Huxley-style kinetics with an integration time step of 250 μs . Current (I) from each channel type was given by $I = g_a b(v - E)$, where g is the local conductance density, a is an activation variable with x order kinetics, b is an optional inactivation variable, v is the local membrane potential, and E is the reversal potential for the ionic species ($E_{\text{Na}} = -70$ mV, $E_{\text{K}} = -90$ mV, $E_{\text{Na}} = 50$ mV, $E_{\text{Ca}} = 140$ mV). Internal calcium concentration was computed using entry via I_{Ca} and removal by a first order pump: $d[\text{Ca}^{2+}]_i/dt = (-1 \times 10^{-5} \times I_{\text{Ca}}/2F) - ([\text{Ca}^{2+}]_i - [\text{Ca}^{2+}]_\infty)/\tau_R$, where $[\text{Ca}^{2+}]_\infty = 0.1 \mu\text{M}$, and $\tau_R = 200$ ms. Channel activation and inactivation variables were expressed in terms of a steady state value, $a_\infty(v)$, and a time constant $\tau_a(v)$ which were calculated from a first-order reaction scheme with forward rate α and backward rate β , giving $a_\infty(v) = \alpha/(\alpha + \beta)$, $\tau_a = 1/(\alpha + \beta)$. The specific rate functions for each current were I_{Na} activation ($x = 3$): $\alpha = 0.182(v + 25)/(1 - e^{-(v+25)/9})$, $\beta = -0.124(v + 25)/(1 - e^{-(v+25)/9})$; I_{Na} inactivation: $\alpha = 0.024(v + 40)/(1 - e^{-(v+40)/5})$, $\beta = -0.0091(v + 65)/(1 - e^{-(v+65)/5})$; I_{Ca} activation ($x = 2$): $\alpha = 0.055(v + 27)/(1 - e^{-(v+27)/3.8})$, $\beta = 0.94e^{-(v+75)/17}$; I_{Ca} inactivation: $\alpha = 4.57 \times 10^{-4}e^{-(v+13)/50}$, $\beta = 0.0065/(1 + e^{-(v+15)/28})$; I_{K} activation ($x = 1$): $\alpha = 0.02(v - 25)/(1 - e^{-(v-25)/9})$, $\beta = -0.002(v - 25)/(1 - e^{-(v-25)/9})$; I_{Ks} activation ($x = 1$): $\alpha = 1 \times 10^{-4}(v + 30)/(1 - e^{-(v+30)/9})$, $\beta = -1.10^{-4}(v + 30)/(1 - e^{-(v+30)/9})$; I_{KCa} activation ($x = 1$): $\alpha([\text{Ca}^{2+}]_i) = 0.01 \times [\text{Ca}^{2+}]_i$, $\beta = 0.02$. Specific membrane capacitance (C_m) was



0.75 $\mu\text{F cm}^{-2}$ (except myelinated axon segments, where $C_m = 0.02 \mu\text{F cm}^{-2}$). Specific membrane resistance (R_m) was 30 $\text{k}\Omega \cdot \text{cm}^2$ (except axon node segments, where $R_m = 50 \Omega \cdot \text{cm}^2$). Specific axial resistance was 150 $\Omega \cdot \text{cm}$. Conductance densities (in $\text{pS } \mu\text{m}^{-2}$) were as follows. Dendrites: $g_{\text{Na}} = 20$, $g_{\text{Ca}} = 0.3$, $g_{\text{KCa}} = 3$, and $g_{\text{Kv}} = 0.1$. Soma: as dendrites and in addition $g_{\text{Kv}} = 200$. Axon hillock and initial segment: $g_{\text{Kv}} = 2000$, $g_{\text{Na}} = 30,000$. Nodes of Ranvier: $g_{\text{Na}} = 30,000$. The rates and conductance densities were developed at 23 $^{\circ}\text{C}$ and were therefore increased from the values given to 37 $^{\circ}\text{C}$ using a Q_{10} of 2.3.

FIG. 2 Effects of electrical structure on firing pattern in a reduced model. **a**, A two-compartment model incorporating the same channels modelled in Fig. 1. The two compartments correspond to the dendritic tree ('dendrite') and the soma and axon initial segment ('axon-soma'). The parameter κ specifies the electrical resistance (coupling) between the two compartments. The parameter ρ specifies the ratio of dendritic to axo-somatic area and thereby sets the strength of dendritic currents relative to axo-somatic currents. The channels and membrane properties of each compartment are depicted. **b**, Uncoupled dendritic (top) and axo-somatic (bottom) compartments ($\kappa \rightarrow \infty$) are each capable of discharging repetitively when current is injected (top, 400 pA; bottom, 10 pA). Note that the firing frequency of the axo-somatic compartment, which is driven by the fast I_{Na} and I_{Kv} is much higher than for the dendritic compartment, which is driven by I_{Ca} , I_{KCa} and I_{Km} . **c**, When fully coupled ($\kappa \rightarrow 0$), eliminating electrotonic effects, the amount of spike-frequency adaptation varies with the size of the dendritic compartment, but the model does not display bursting or spike ADPs. **d**, **e**, Partial coupling produces voltage gradients between axo-somatic (left) and dendritic (right) compartments and supports bursting and ADPs. When partially coupled, changes in either dendritic area ρ (**d**) or changes in coupling κ (**e**) alter firing pattern to injected current (100 pA, injected in axo-somatic compartment). Scale bar (30 mV, 200 ms) applies to all panels. **METHODS**. Compartmental simulations were carried out with channels described in Fig. 1. The dendritic compartment properties were: $C_m = 0.75 \mu\text{F cm}^{-2}$, $R_m = 30 \text{ k}\Omega \text{ cm}^2$, and active conductances (in $\text{pS } \mu\text{m}^{-2}$) $g_{Na} = 15$, $g_{Ca} = 0.3$, $g_{KCa} = 3$, and $g_{Km} = 0.1$. The axo-somatic compartment



ment contained just $g_{Kv} = 1500$, $g_{Na} = 30,000$. The compartments were connected with axial resistance given by the parameter κ , which generally ranged from 1 to 10 $\text{M}\Omega$. The area of the axo-somatic compartment was $100 \mu\text{m}^2$ and the area of the dendritic compartment was specified as a multiple (ρ) of the axo-somatic area, generally 100 to 500. Presence of a leak conductance and capacitance in the axo-somatic compartment did not affect the results and were therefore omitted for simplicity.

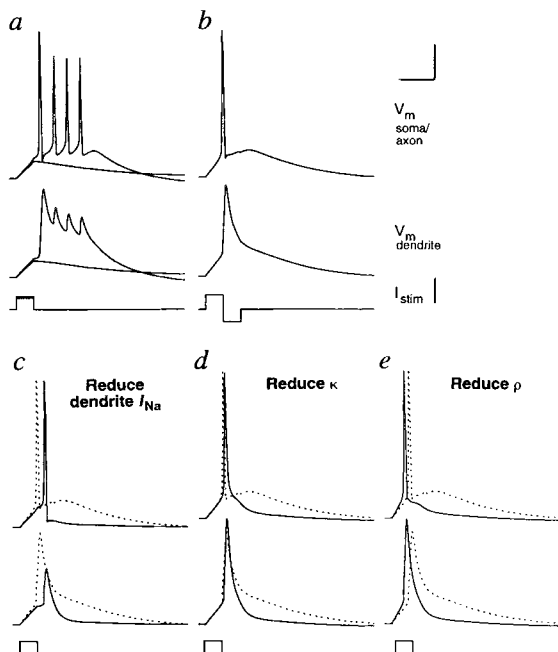


FIG. 3 Electrical basis for spike after-depolarization and bursting in the reduced model. **a**, All-or-none bursting in a partially coupled two-compartment model ($\rho = 200$, $\kappa = 10$) was triggered by a short current pulse. Just subthreshold and superthreshold responses are shown. **b**, The burst could be suppressed after the first spike when followed by a short hyperpolarizing pulse. This revealed an underlying depolarizing envelope driven by a prolonged dendritic spike. **c**, Reducing the dendritic Na^+ conductance by 70% reduced the width and amplitude of the dendritic spike and suppressed burst generation and the depolarizing potential following the spike. Reducing the electrical coupling κ (**d**, 50% reduction) or the ratio of dendritic to axo-somatic area ρ (**e**, 30% reduction) also substantially reduced the sustained depolarization. In each panel, voltage of the axo-soma (top) and the dendrite (centre) and the stimulus (bottom) are shown. In panels **c**–**e**, the control condition **b** is shown in dotted lines for comparison. Scale bars (20 ms, 30 mV, 1 nA) apply to all panels.

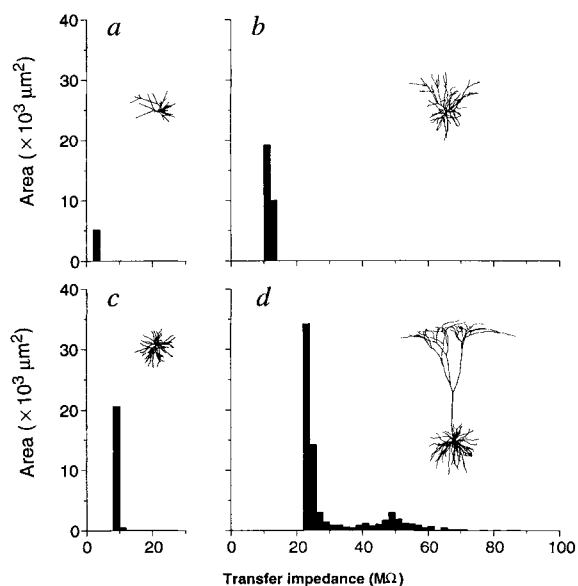


FIG. 4 Electrical geometry of cortical cells. Histograms show the distribution of electrical attenuation between the soma and dendritic segments for the four neurons depicted in Fig. 1. The steady-state transfer impedance (Z) from the soma to each simulated dendritic compartment was calculated by injecting a small current step (I) in the soma and measuring the resulting (passive) steady-state voltage change (V) in the dendritic compartment ($Z = V/I$). The histogram bin corresponding to this impedance level was then incremented by an amount equal to the membrane area of that compartment. The total area of the histogram therefore reflects the total dendritic area and the shape of the histogram reflects the relative electrical distance of the dendritic membrane from the soma. These can be compared to the parameters ρ and κ in the reduced model, respectively.

Direct experimental tests of this model could be made by partial dendritic ablation or manipulations to alter cytoplasmic resistivity.

As there is a wide anatomical variety of neocortical dendrites¹, our findings support the idea of a continuous spectrum of neocortical firing patterns^{2,3} rather than discrete categories. Differences in spike shape^{2,24} passive response properties²⁴, and some aspects of the time course of adaptation²⁵ observed *in vitro* were not readily reproduced in the present model. These limitations indicate that differences in channel types or intracellular Ca^{2+} dynamics may also be important. Thus, although the present study does not exclude contributions of differential channel expression or other physiological differences, it supports the hypothesis that neocortical neurons that share similar channel distributions may derive functional differentiation from their dendritic morphology. Similar principles may also apply to other morphologically heterogeneous neuronal populations. □

Received 10 April; accepted 7 June 1996.

- Peters, A. & Jones, E. G. *Cerebral Cortex Vol 1: Cellular Components of the Cerebral Cortex* (Plenum, New York, 1984).
- McCormick, D. A., Connors, B. W., Lighthall, J. W. & Prince, D. A. *J. Neurophysiol.* **54**, 782–806 (1985).
- Connors, B. W. & Gutnick, M. J. *Trends Neurosci.* **13**, 99–104 (1990).
- Chagnac-Amitai, Y., Luhmann, H. J. & Prince, D. A. *J. comp. Neurol.* **296**, 598–613 (1990).
- Mason, A. & Larkman, A. U. *J. Neurosci.* **10**, 1415–1428 (1990).
- Franceschetti, S., Gualtiero, E., Panzica, F., Sancini, G. E. W. & Avanzini, G. *Brain Res.* **696**, 127–139 (1995).
- Yang, C. R., Seamans, J. K. & Gorelova, N. J. *Neurosci.* **16**, 1904–1921 (1996).
- Mainen, Z. F., Joerges, J., Huguenard, J. R. & Sejnowski, T. J. *Neuron* **15**, 1427–1439 (1995).
- Yuste, R. & Tank, D. W. *Neuron* **16**, 701–716 (1996).
- Rapp, M., Yarom, Y. & Segev, I. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Stuart, G. J. & Sakmann, B. *Nature* **367**, 69–72 (1994).
- Wollner, D. A. & Catterall, W. A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8424–8428 (1986).
- Angelides, K. J., Elmer, L. W., Loftus, D. & Elson, E. J. *Cell Biol.* **106**, 1911–1924 (1988).
- Schwindt, P. C. et al. *J. Neurophysiol.* **59**, 424–449 (1988).
- Storm, J. F. *Prog. Brain Res.* **83**, 161–187 (1990).
- Kim, H. G. & Connors, B. W. *J. Neurosci.* **13**, 5301–5311 (1993).
- Yuste, R., Gutnick, M. J., Saar, D., Delaney, R. D. & Tank, D. W. *Neuron* **13**, 23–43 (1994).
- Pinsky, P. F. & Rinzel, J. *J. comput. Neurosci.* **1**, 39–60 (1994).

- Turner, R. W., Maler, L., Deerinck, T., Levinson, S. R. & Ellisman, M. H. *J. Neurosci.* **14**, 6453–6471 (1994).
- Azouz, R., Jensen, M. S. & Yaari, Y. *J. Physiol.* **492**, 211–223 (1996).
- Granit, R., Kernell, D. & Smith, R. S. *J. Physiol.* **168**, 100–115 (1963).
- Kandel, E. R. & Spencer, W. A. *J. Neurophysiol.* **24**, 243–259 (1961).
- Larkman, A. U., Major, G., Stratford, K. J. & Jack, J. B. *J. comp. Neurol.* **323**, 137–152 (1992).
- Kasper, E. M., Larkman, A. U., Lubke, J. & Blakemore, C. J. *comp. Neurol.* **339**, 475–494 (1994).
- Agmon, A. & Connors, B. W. *J. Neurosci.* **12**, 319–329 (1993).
- Hines, M. In *Neural Systems: Analysis and Modeling* (ed. Eeckman, F. H.) 127–136 (Kluwer, Boston, MA, 1993).
- Hamill, O. P., Huguenard, J. R. & Prince, D. A. *Cerebral Cortex* **1**, 48–61 (1991).
- Gutfreund, Y., Yarom, Y. & Segev, I. *J. Physiol.* **483**, 621–640 (1995).
- Reuveni, I., Friedman, A., Amitai, Y. & Gutnick, M. J. *J. Neurosci.* **13**, 4609–4621 (1993).
- Sloper, J. J. & Powell, T. P. S. *Phil. Trans. R. Soc. Lond.* **B285**, 173–197 (1978).

ACKNOWLEDGEMENTS. Z.F.M. was supported by an H.H.M.I. predoctoral fellowship. T.J.S. is supported by the H.H.M.I., the N.I.M.H. and the O.N.R. We thank D. K. Smetters, R. Douglas, K. Martin, B. Connors, L. Caulier and J. Anderson for dendritic reconstructions; J. Huguenard, J. Rinzel and P. Rhodes for helpful discussions; and A. Destexhe, G. Brown and R. Ritz for comments on the manuscript.

CORRESPONDENCE and requests for materials should be addressed to Z.F.M. (e-mail: zach@salk.edu). NEURON code for these simulations is available electronically at <http://www.cnl.salk.edu/CNL/simulations.html>.

Critical role for $\beta 7$ integrins in formation of the gut-associated lymphoid tissue

Norbert Wagner*, Jürgen Löhler†, Eric J. Kunkel‡, Klaus Ley‡, Euphemia Leung§, Geoff Krissansens§, Klaus Rajewsky* & Werner Müller*

* Institute for Genetics, University of Cologne, 50937 Cologne, Germany

† Heinrich-Pette-Institute for Experimental Virology and Immunology, University of Hamburg, 20251 Hamburg, Germany

‡ Department of Biomedical Engineering, University of Virginia, Charlottesville, Virginia 22908, USA

§ Department of Molecular Medicine, University of Auckland, Auckland, New Zealand

IMMUNE defence against pathogens entering the gut is accomplished by lymphocytes in the gut-associated lymphoid tissue (GALT), a major compartment of the immune system¹. The GALT, comprising Peyer's patches, lamina propria lymphocytes and intra-epithelial lymphocytes of the intestine, is populated by lymphocytes that migrate there from the vasculature^{2,3}. Here we report that, in mice deficient for the $\beta 7$ integrin subfamily of adhesion molecules, the formation of the GALT is severely impaired. This is probably due to a failure of $\beta 7^{-/-}$ lymphocytes to arrest and adhere to the vasculature at the site of transmigration into the GALT.

Lymphocytes move through the GALT by means of differential expression and activation of adhesion molecules⁴. Emigration of lymphocytes from the vasculature into tissues requires them to roll along the endothelial cells, arrest and adhere firmly to endothelial cells, and migrate across the endothelium^{5,6}.

The $\beta 7$ integrin subfamily of adhesion molecules is expressed on a subset of lymphocytes, and consists of two members, $\alpha 4\beta 7$ and $\alpha E\beta 7$ (refs 7–10): $\alpha 4\beta 7$ mediates the adherence of lymphocytes to high endothelial venules (HEVs) of Peyer's patches¹¹ and has been implicated in lymphocyte rolling¹², and $\alpha E\beta 7$ mediates the adherence of intra-epithelial lymphocytes of the intestine (IELs) to the intestinal epithelium¹³. In chronic inflammatory bowel disease, either $\alpha 4\beta 7$ or $\alpha 4\beta 1$, or both, mediates the migration of lymphocytes into inflammatory lesions¹⁴. To analyse the contribution of the $\beta 7$ integrins to the formation of the GALT, and to determine how $\beta 7$ integrins participate in the compartmentalization of the immune system, we generated $\beta 7$ -deficient mice.

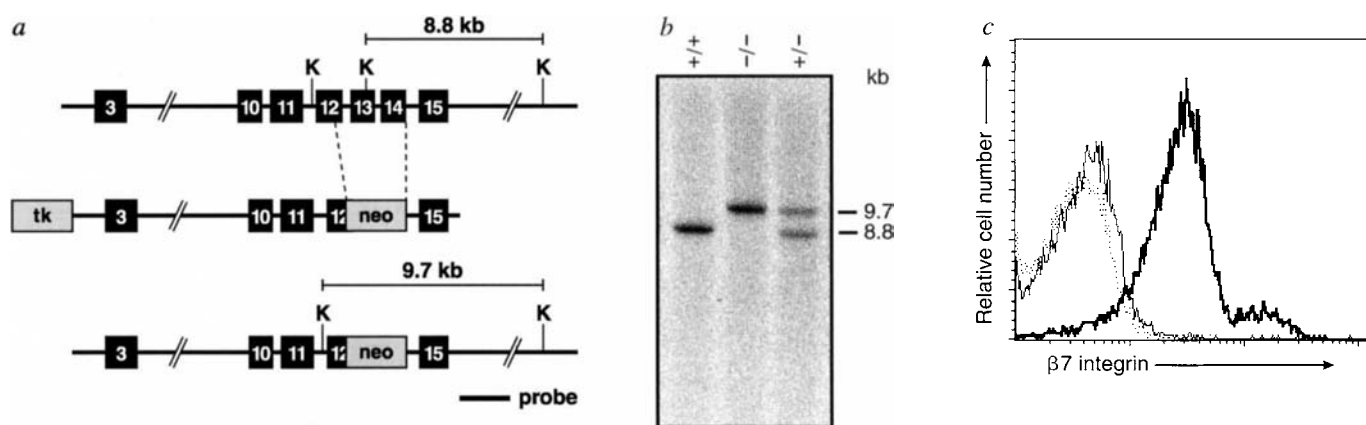


FIG. 1a, Targeted disruption of the gene for $\beta 7$ integrin in mice. Structure of the gene encoding murine $\beta 7$ (top), the targeting vector (middle) and the predicted targeted $\beta 7$ allele (bottom). Numbered boxes denote exons. The lengths of diagnostic restriction fragments and location of the probe 3' of the sequence used for the targeting vector are shown. The neomycin (neo) expression cassette replaces part of exon 12 and both exons 13 and 14, the latter encoding the transmembrane region of the gene for $\beta 7$ integrin. K, KpnI. b, Southern blot analysis of mice derived from intercrossing mice heterozygous for the $\beta 7$ mutation. Homozygous mutant mice ($-/-$) only show presence of the 9.7-kb fragment. c, Flow cytometric analysis of peripheral blood lymphocytes (PBLs) for expression of $\beta 7$. Staining of PBLs of $\beta 7^{+/+}$ mice (thick line) and $\beta 7^{-/-}$ mice (thin line) with the M293 antibody is shown. The broken line represents the negative control. METHODS. Two overlapping murine $\beta 7$ integrin clones (*muR4*, *18L*), spanning from exon 3 to the 3' untranslated region, were combined using a common *XbaI* site. A fragment extending from exon 3 to the *MscI* site in

exon 12 was flanked by an HSV-TK expression cassette at the 5' end, and by a neo expression cassette at the 3' end. A 750-bp fragment containing exon 15 and the following intronic region was generated by PCR and inserted 3' to the neo expression cassette. After linearization by *Clal*, B6III and E14-1 embryonic stem cells were transfected, homologous recombinants were identified by Southern blotting and two independent $\beta 7^{-/-}$ mouse lines were established following previous protocols^{25,26}. The two mouse lines showed a similar phenotype with respect to the impaired formation of the GALT. For the experiments shown here we used the $\beta 7^{-/-}$ C57BL/6 mouse line. Single colour immunofluorescence of PBLs or spleen cells stained with anti- $\beta 7$ monoclonal antibody (mAb) (M293, M298, M300, M301, FIB21), and the DATK32 mAb reactive with a combinatorial epitope of $\alpha 4\beta 7$ (ref. 27), was analysed with a FACScan (Becton-Dickinson). The negative control represents the fluorescence intensity without addition of antibodies.

The $\beta 7$ gene was disrupted by replacing the sequence coding for the transmembrane and part of the extracellular domains with a neo expression cassette (Fig. 1a). We identified $\beta 7^{+/+}$, $\beta 7^{-/-}$, and $\beta 7^{-/-}$ offspring by Southern blotting (Fig. 1b), but expression of $\beta 7$ and $\alpha 4\beta 7$ was not detected on lymphocytes in $\beta 7^{-/-}$ mice (Fig. 1c).

The $\beta 7^{-/-}$ mice exhibited an overall hypoplasia of the GALT. No Peyer's patches were macroscopically detectable at the age of six weeks, whereas ~8–12 Peyer's patches were found to protrude out of the small intestine in $\beta 7^{+/+}$ mice. Microscopically, the average diameter of the Peyer's patches in $\beta 7^{-/-}$ mice was reduced to 10–20% of that in $\beta 7^{+/+}$ mice, despite $\beta 7^{-/-}$ mice having an increased load of pathogens in the small intestine as compared to wild-type mice (Fig. 2a, and data not shown). These small Peyer's patches with decreased cellularity and rudimentary follicles did not disrupt the contours of the intestinal wall (Fig. 2a, b). However, when the number of Peyer's patches in $\beta 7^{-/-}$ and wild-type mice was determined in serial sections there was no numerical difference (8–12) between the two groups, suggesting that the stromal elements of the anlage of the Peyer's patches were not disturbed.

The intestine is an important site for production of secretory IgA¹. By analysing sections of lamina propria we estimate that there was a 10–30-fold difference in the number of IgA⁺ and IgM⁺ B cells and plasma cells between $\beta 7^{-/-}$ mice and wild-type mice (Fig. 2c, and data not shown). A similar reduction was noted for IgM⁺ cells in Peyer's patches and CD4⁺ cells in the lamina propria of $\beta 7^{-/-}$ mice (Fig. 2b, d). The IELs located between the epithelial cells were also reduced in number in $\beta 7^{-/-}$ mice; the mucosal IEL index, calculated from the ratio of CD8⁺ T cells to epithelial cells, was 3.3 in the mutants versus 17 in control mice ($P < 0.005$; Fig. 2e)¹⁵.

The GALT was compromised at all sites in $\beta 7^{-/-}$ mice. However, this was not caused by defective lymphocyte development, as flow cytometric analysis of lymphocytes in bone marrow, thymus,

spleen, peripheral lymph nodes (PLNs) and mesenteric lymph nodes (MLNs) did not indicate any differences between mutant and wild-type mice (data not shown). The stromal elements required for the formation of the GALT seem to be present and functional in $\beta 7^{-/-}$ mice because transfer of bone marrow from $\beta 7^{+/+}$ donors into irradiated $\beta 7^{-/-}$ hosts led to reconstitution of the GALT with $\beta 7^{+/+}$ lymphocytes (not shown).

The composition of the reduced lymphocyte population in the Peyer's patches of mutant mice was similar to that of wild-type mice when assessed for B- and T-lineage markers. Expression of the $\alpha \beta$ and $\gamma \delta$ T-cell receptors was found on some IELs in mutant mice (not shown). Lymphocytic infiltrates were noted in the lungs of $\beta 7^{-/-}$ mice, indicating that migration of lymphocytes to pulmonary sites is not exclusively dependent on $\beta 7$ integrins (not shown).

Is the impaired GALT formation in $\beta 7^{-/-}$ mice caused by altered binding of $\beta 7^{-/-}$ lymphocytes to HEVs in Peyer's patches? In the HEV-binding assay¹⁶, the adherence of $\beta 7^{-/-}$ lymphocytes to HEVs of wild-type Peyer's patches was almost abrogated, and the number of bound $\beta 7^{-/-}$ lymphocytes to HEVs of wild-type MLNs was reduced by 45% compared to the number of bound $\beta 7^{+/+}$ lymphocytes ($P < 0.005$; Fig. 3a). To assess the *in vivo* effects of the $\beta 7$ mutation on lymphocyte migration into lymphoid tissue, short-term migration assays were performed¹⁷. The migration of $\beta 7^{-/-}$ lymphocytes into Peyer's patches was almost abolished, and migration into MLNs was reduced by 49% ($P < 0.005$ and $P < 0.05$, respectively; Fig. 3b). Therefore, lymphocyte migration to Peyer's patches is dependent on $\beta 7$, but migration into MLNs is partly mediated by adhesion molecules other than $\beta 7$.

To determine whether the defective migration of $\beta 7^{-/-}$ lymphocytes *in vivo* resulted from a defect in leukocyte rolling or in arrest and firm adhesion to endothelial cells, we used epifluorescence videomicroscopy in Peyer's patches. No difference was observed in leukocyte-rolling flux¹⁸ in $\beta 7^{-/-}$ and $\beta 7^{+/+}$ mice

(Fig. 4a). However, the average rolling velocity in $\beta 7^{-/-}$ mice was slightly higher than in $\beta 7^{+/+}$ mice ($88 \pm 9 \mu\text{m s}^{-1}$ versus $62 \pm 7 \mu\text{m s}^{-1}$; $P < 0.05$), indicating that $\beta 7$ may slow down lymphocyte rolling. Strikingly, rolling leukocytes in $\beta 7^{-/-}$ mice exhibited a major defect in arrest and firm adhesion to the endothelial cells, as representatively shown on a videotape of HEVs in Peyer's patches ($P < 0.005$, Fig. 4b, c). Therefore, we believe that the absence of $\beta 7$ causes a defect in the transition from lymphocyte rolling to firm adhesion in the HEVs of Peyer's patches.

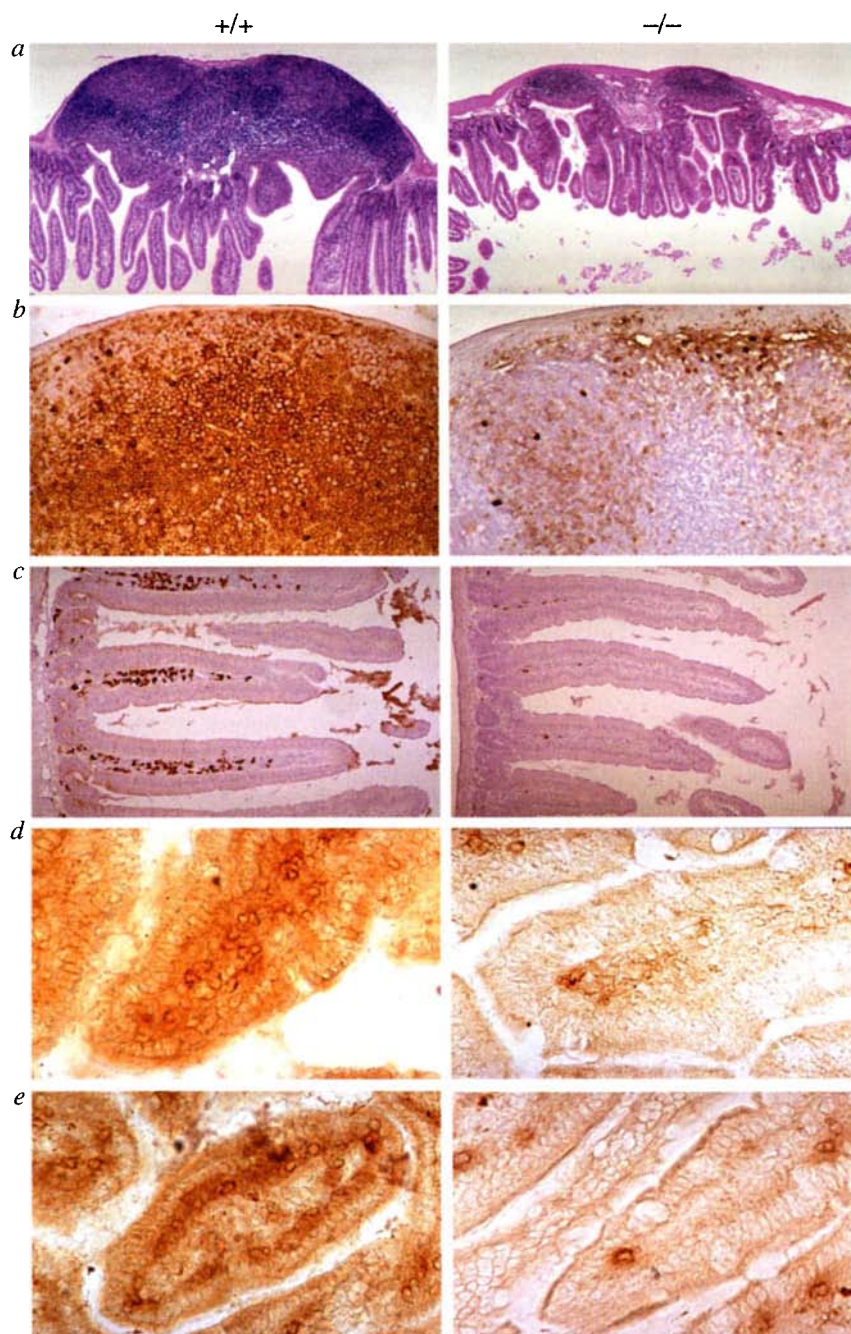
Taken together, our findings indicate that the formation of the GALT depends on $\beta 7$ -mediated adhesive interactions between lymphocytes and endothelial cells. Of the known members of the $\beta 7$ integrin subfamily, $\alpha 4\beta 7$ should be responsible for lymphocyte migration to Peyer's patches and lamina propria, as $\alpha \text{E}\beta 7$ is expressed on only a few lymphocytes other than IELs¹⁹. Lymphocytes found in Peyer's patches and lamina propria of $\beta 7^{-/-}$ mice

might have used adhesion molecules other than $\alpha 4\beta 7$ for inefficient transmigration through the vasculature, or they might have migrated by routes other than transmigration (for example, lymphatic vessels connecting intestinal villi with Peyer's patches)²⁰; the residual IELs might have developed *in situ*^{21,22}.

Our results help to define the difference in function of $\alpha 4\beta 7$ as compared to $\alpha \text{E}\beta 7$ and $\alpha 4\beta 1$. Haematopoietic stem cells did not colonize fetal liver in chimaeric mice lacking the $\beta 1$ integrins²³. A defective migration of pro-thymocytes from bone marrow to thymus has been demonstrated in $\alpha 4$ integrin-deficient chimaeric mice²⁴, and $\alpha 4^{-/-}$ lymphocytes were able to colonize the intra-epithelial sites of the intestine²⁴. In $\beta 7^{-/-}$ mice, lymphocyte development was normal and the number of IELs was reduced. Therefore, it seems likely that $\alpha 4\beta 1$ mediates migration of pro-thymocytes from bone marrow to thymus, and $\alpha \text{E}\beta 7$ is probably responsible for the immobilization of a large proportion of the IELs within the intestinal epithelium.

FIG. 2 Histological and immunohistochemical analysis of the GALT in $\beta 7^{+/+}$ (left column) and $\beta 7^{-/-}$ mice (right column). a, The $\beta 7^{+/+}$ Peyer's patch is composed of three individual follicles protruding into the enteric lumen and the peritoneal cavity, with large germinal centres and prominent follicle-associated epithelium. The $\beta 7^{-/-}$ hypoplastic Peyer's patch consists of two small, rudimentary follicles less densely populated by lymphocytes and discrete follicle-associated epithelium. Note the disproportionate dimensions of the marginally localized postcapillary venules and lymph vessels in $\beta 7^{-/-}$. b, Immunostaining for IgM⁺ B cells and plasma cells (brown) on paraffin sections of Peyer's patches. c, Immunostaining for IgA⁺ B cells and plasma cells on paraffin sections of the duodenal lamina propria. d, Immunostaining for CD4⁺ cells in the jejunal lamina propria from frozen tissue sections. e, Immunostaining for CD8⁺ IELs in the jejunum from frozen tissue sections. Magnification: a, $\times 16$; b, $\times 60$; c, $\times 40$; d, e, $\times 630$. Most $\beta 7^{-/-}$ mice exhibited mild inflammatory changes of the small bowel, consisting of neutrophilic and eosinophilic granulocytes as well as macrophages infiltrating the lamina propria (not shown). This could reflect a compensatory mechanism to cope with the increased load of micro organisms in the intestines of $\beta 7^{-/-}$ mice.

METHODS. Intestines were dissected and intestinal rolls were prepared as described¹⁸. The paraffin blocks were serially sectioned, and every fifth section was stained alternatively with haematoxylin and eosin and periodic-acid Schiff reaction, and Peyer's patches were counted microscopically. Immunostaining of IgM and IgA used the avidin-biotin-peroxidase complex method²⁸. Rabbit anti-mouse α and μ heavy-chain antibodies were from Nordic (Tilburg, The Netherlands); biotin-conjugated F(ab')₂ fragments of sheep anti-rabbit IgG were from Sigma; Germany; and streptavidin-biotinylated peroxidase complexes were from DAKO Diagnostika (Hamburg). For immunostaining of frozen sections, sections were fixed in acetone, incubated with 0.1% phenylhydrazine (Sigma) in PBS and blocked with PBS containing 4% of fetal calf serum. For staining, biotinylated mAb GK1.5 (anti-CD4) and 53-6.7 (anti-CD8) were used. Primary antibodies were detected with peroxidase-conjugated streptavidin (Dianova, Hamburg).



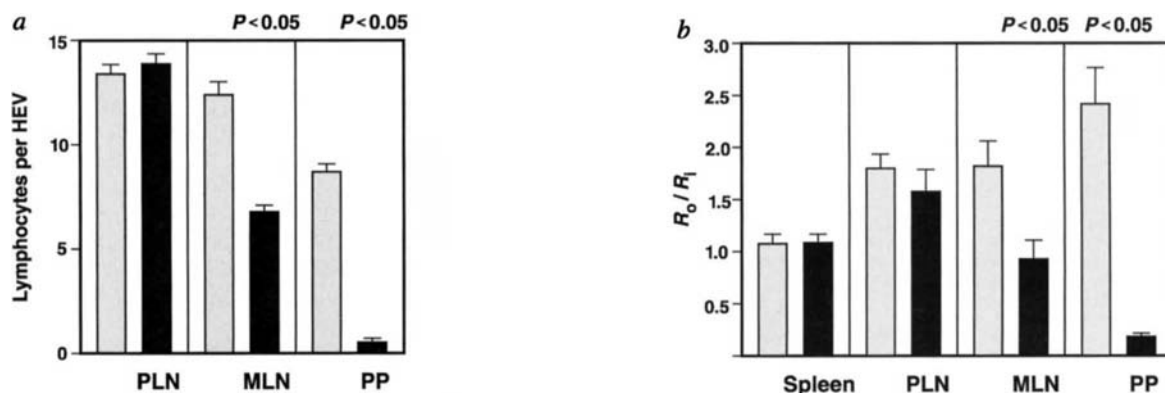


FIG. 3 Impaired adherence and migration of $\beta 7^{-/-}$ lymphocytes to the GALT. **a**, Binding of spleen lymphocytes isolated from $\beta 7^{+/+}$ (grey bars) and $\beta 7^{-/-}$ (black bars) mice to HEVs from $\beta 7^{+/+}$ PLN, MLN and Peyer's patches (PP). Results were obtained from at least two separate experiments for each lymphoid organ; values are mean number of cells adherent to HEV ($\pm$ s.e.m.) (100 HEVs were evaluated). **b**, Short-term migration of spleen lymphocytes isolated from $\beta 7^{+/+}$ (grey bars) and $\beta 7^{-/-}$ (black bars) mice to spleen, PLN, MLN and Peyer's patches after intravenous injection of labelled single-cell preparations into the lateral tail vein of $\beta 7^{-/-}$ mice. Results represent the mean single-cell preparations into the lateral tail vein of $\beta 7^{+/+}$ mice. Results represent the mean R_0/R_i (see Methods) ($\pm$ s.e.m.) from four mice for each experimental setting.

METHODS. The modified Stamper–Woodruff frozen-section assay with PLN, MLN and Peyer's patches from C57BL/6 mice was used to evaluate lymphocyte adherence to endothelial cells under shear stress¹⁶. Briefly, murine splenocytes were overlaid onto 12- μ m frozen tissue sections and incubated under rotation (70 r.p.m.) for 30 min at 7 °C. After fixation the slides were counterstained with Gill's haematoxylin. We counted the lymphocytes bound per HEV for several slides for each cell sample. For the short-term migration assay¹⁷, single-cell suspensions of spleen were labelled with either FITC or PKH26 (Sigma). FITC-labelled lymphocytes (control or test populations) were mixed with an equal number of an internal standard population of PKH26-labelled lymphocytes and injected i.v. into C57BL/6 mice. After 90 min the ratio of FITC-PKH26 labelled cells (R_0) was determined by flow cytometry in single-cell suspensions prepared from spleen, PLN, MLN and Peyer's patches. The result was corrected for the input ratio of injected cells (R_i). Statistical significance of results from both assays was determined using the unpaired, two-tailed Student's *t*-test.

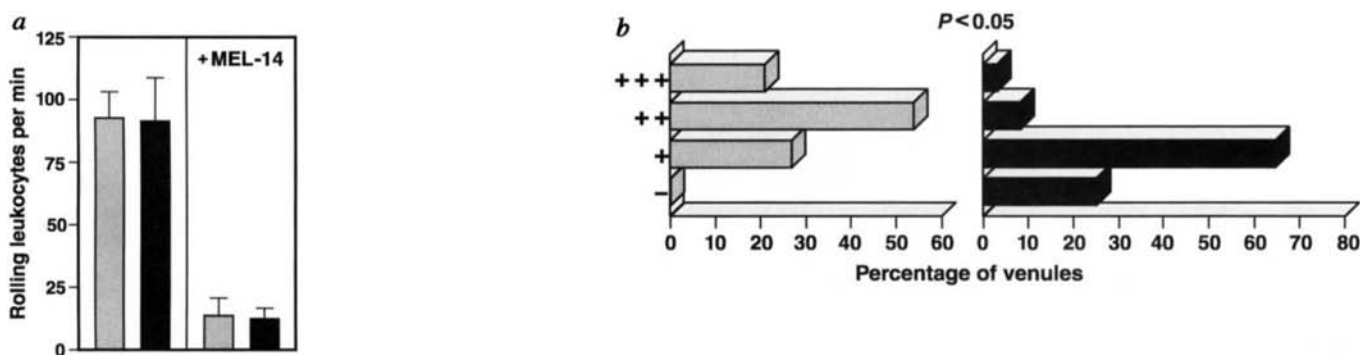


FIG. 4 Intravital epifluorescence microscopy in $\beta 7^{+/+}$ and $\beta 7^{-/-}$ mice¹⁴. **a**, Rolling leukocyte flux in $\beta 7^{+/+}$ (5 mice, 28 venules; grey bars) and $\beta 7^{-/-}$ mice (5 mice, 17 venules; black bars) before (left) and after (right) addition of the anti-L-selectin mAb MEL-14 ($\beta 7^{+/+}$, 2 mice, 7 venules; $\beta 7^{-/-}$ mice, 2 mice, 6 venules; mean $\pm$ s.e.m.). Systemic counts and leukocyte differentials were similar in $\beta 7^{+/+}$ and $\beta 7^{-/-}$ mice. The injection of the MEL-14 mAb reduced rolling in $\beta 7^{-/-}$ and $\beta 7^{+/+}$ by $\sim 90\%$ ($P < 0.005$), suggesting that L-selectin is necessary for lymphocyte rolling in HEVs in Peyer's patches *in vivo*. **b**, Percentage of venules recorded without (–), with some (+), with many (++), and covered with (+++) lymphocytes in $\beta 7^{+/+}$ (grey) and $\beta 7^{-/-}$ mice (black). **c**, Video frame of HEVs from a $\beta 7^{+/+}$ (left) and a $\beta 7^{-/-}$ mouse (right) 30 min after injection of fluorescent dye. Many adherent cells (white spots) can be seen in the $\beta 7^{+/+}$ HEV (diameter, 35.4 μ m), compared to the $\beta 7^{-/-}$ HEV (diameter, 28.4 μ m), with little adhesion and no emigration.

METHODS. Male mice (11) ~ 100 days in age and weighing between 25 and 35 g were used for surgery²⁹. The Peyer's patch was prepared for intravital microscopy as described¹². Leukocytes that had been labelled with acridine red (0.15 mg, i.v.; Chroma, Stuttgart) were observed for adhesive interactions with HEVs in Peyer's patches, and recorded using an experimental setup (intravital microscope modified for stroboscopic epifluorescence illumination, video recorder and camera system) as described²⁹. Microvessel diameters, the rolling leukocyte flux (in cells per min), centre-line blood-flow velocity, and leukocyte-rolling velocities were determined as outlined²⁹. Monoclonal antibody MEL-14 (rat IgG1, 100 μ g per mouse) binds and blocks murine L-selectin³⁰. Statistical analysis was done using two-tailed Student's *t*-test for comparison between groups, and Kolmogorov–Smirnov Z-test to see whether the histograms came from identical distributions.

As well as providing a means of determining the role of $\beta 7$ integrins in the compartmentalization of the immune system, the $\beta 7^{-/-}$ mice could be used to study the function of the GALT in the development of oral tolerance and the pathogenesis of chronic inflammatory bowel disease. □

Received 26 April; accepted 14 June 1996.

1. Brandtzaeg, P. *Curr. Topics Microbiol. Immun.* **146**, 13–25 (1989).
2. Butcher, E. C. & Picker, L. J. *Science* **272**, 60–66 (1996).
3. Imhof, B. A. & Dunon, D. *Adv. Immun.* **58**, 345–416 (1995).
4. Carlos, T. M. & Harlan, J. M. *Blood* **84**, 2068–2101 (1994).
5. Butcher, E. C. *Cell* **67**, 1033–1036 (1991).
6. Springer, T. A. *Cell* **76**, 301–314 (1994).
7. Holzmann, B., McIntyre, B. W. & Weissman, I. L. *Cell* **56**, 37–46 (1989).
8. Holzmann, B. & Weissman, I. L. *EMBO J.* **8**, 1735–1741 (1989).
9. Cerf-Bensussan, N., Begue, B., Gagnon, J. & Meo, T. *Eur. J. Immun.* **22**, 273–277 (1992).
10. Parker, C. M. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 1924–1928 (1992).
11. Hamann, A., Andrew, D. P., Jablonski-Westrich, D., Holzmann, B. & Butcher, E. C. *J. Immun.* **152**, 3282–3293 (1994).
12. Bargatze, R. F., Jutila, M. A. & Butcher, E. C. *Immunology* **3**, 99–108 (1995).
13. Cepek, K. L. et al. *Nature* **372**, 190–193 (1994).
14. Podolsky, D. K. et al. *J. clin. Invest.* **92**, 372–380 (1993).
15. Rocha, B., Vassalli, P. & Guy-Grand, D. *J. exp. Med.* **180**, 681 (1994).
16. Butcher, E., Scollay, R. & Weissman, I. J. *Immun.* **123**, 1996–2003 (1979).
17. Arbones, M. L. et al. *Immunology* **1**, 247–260 (1994).
18. Sundberg, J. P., Elson, C. O., Bedigian, H. & Birkenmeier, E. H. *Gastroenterology* **107**, 1726–1735 (1994).
19. Kilshaw, P. J. & Murrant, S. J. *Eur. J. Immun.* **21**, 2591–2597 (1991).
20. Griebel, P. J. & Hein, W. R. *Immun. Today* **17**, 30–39 (1996).
21. Guy-Grand, D. et al. *Eur. J. Immun.* **22**, 505–510 (1992).
22. Sim, G. K., Rajasekar, R., Dessing, M. & Augustin, A. *In. Immun.* **6**, 1287–1295 (1994).
23. Hirsch, E., Iglesias, A., Potocnik, A. J., Hartmann, U. & Fässler, R. *Nature* **380**, 171–175 (1996).
24. Arroyo, A. G., Yang, J. T., Rayburn, H. & Hynes, R. O. *Cell* **85**, 997–1008 (1996).
25. Kühn, R., Rajewsky, K. & Müller, W. *Science* **254**, 707–710 (1991).
26. Zou, Y. R., Takeda, S. & Rajewsky, K. *EMBO J.* **12**, 811–820 (1993).
27. Andrew, D. et al. *J. Immun.* **153**, 3847 (1994).
28. Hsu, S., Raine, L. & Fanger, H. J. *Histochem. Cytochem.* **29**, 577–580 (1981).
29. Kunkel, E. J. et al. *J. exp. Med.* **183**, 57–65 (1996).
30. Gallatin, W. M., Weissman, I. L. & Butcher, E. C. *Nature* **304**, 30–34 (1983).

ACKNOWLEDGEMENTS. We thank B. Hampel, A. Egert and G. von Hesberg for technical assistance; M. Hue and I. Weissman for the murine $\beta 7$ cDNA; P. J. Kilshaw and W. Newman for monoclonal antibodies against $\beta 7$ and $\alpha 4\beta 7$; U. Ringeisen for artwork; B. Sydora for help with the bone-marrow transfer experiment; and A. Tarakhovskiy and R. Rickert for critical reading of the manuscript. This work was supported by the Bundesministerium für Bildung und Forschung through the Genzentrum Köln. N.W. is a recipient of an AIDS grant of the Deutsches Krebsforschungszentrum and the Bundesministerium für Bildung und Forschung.

CORRESPONDENCE and requests for materials should be addressed to N.W. (e-mail: nwagner@gen1.genetik.uni-koeln.de).

Role of kinases and the phosphatase calcineurin in the nuclear shuttling of transcription factor NF-AT4

Futoshi Shibasaki, E. Roydon Price, David Milan & Frank McKeon

Department of Cell Biology, Harvard Medical School, Boston, Massachusetts 02115, USA

A NEW facet of calcium signalling involves the nuclear import of the NF-AT transcription factors from their dormant position in the cytoplasm^{1–3}. The protein phosphatase calcineurin appears to play an essential role in activating NF-AT nuclear import, as the calcineurin inhibitors cyclosporin A and FK506 block dephosphorylation and nuclear import of NF-AT (refs 4–7). Here we show that calcium signalling induces an association between NF-AT4 and calcineurin, and that these molecules are transported, as a complex, to the nucleus, where calcineurin continues to dephosphorylate NF-AT4. We propose that a nuclear complex of NF-AT4 and calcineurin maintains calcium signalling by counteracting a vigorous nuclear NF-AT kinase.

NF-AT4 (Fig. 1a) is expressed at high levels in CD4⁺CD8⁺ thymocytes^{8,9}. To test the possibility that NF-AT4 could respond

to calcium signals in non-T cells, a green fluorescent protein¹⁰ (GFP double mutant: T65/A163) fused with NF-AT4 was expressed in BHK fibroblasts (Fig. 1b). GFP–NF-AT4 is exclusively cytoplasmic in unstimulated cells and is transported to the nucleus over a 10-min period following exposure to calcium ionophore (Fig. 1b, top panels). These same cells export GFP–NF-AT4 in the 20 min after removal of ionophore (Fig. 1b, middle panels). NF-AT4 can participate in multiple rounds of calcium signalling without detectable changes in kinetics of either import or export (Fig. 1b, bottom panels). Analysis of a larger population of cells revealed a $t_{1/2}$ for NF-AT4 nuclear import and export of ~ 7.5 min and 12 min, respectively (Fig. 1c). These kinetics are unaffected by the presence of cycloheximide, and are therefore a reflection of dynamics rather than *de novo* NF-AT4 synthesis (Fig. 1c).

During the activation of T cells, NF-AT1 is dephosphorylated in a manner detectable by a shift to a higher-mobility form in gel electrophoresis^{6,7,11,12}. We compared the timing of NF-AT4 electrophoretic mobility shifts with that of NF-AT4 nuclear import and export (Fig. 1d). In the presence of calcium ionophore, NF-AT4 undergoes a progressive mobility shift, with a $t_{1/2}$ of ~ 10 min. Notably, the conversion of NF-AT4 to the slower-mobility form upon washing out the calcium ionophore is very rapid, with a $t_{1/2}$ of less than 5 min (Fig. 1d). Calcineurin activity was shown to be essential for both NF-AT4 nuclear import and the NF-AT4 mobility shift, as these are blocked by both cyclosporin A (CsA) and FK506 (Fig. 1e, f). Moreover, the constitutively active calcineurin mutant Δ CnA (refs 13–15) activates both the NF-AT4 nuclear import and mobility shifts in the absence of calcium ionophores (Fig. 1f). In addition to showing calcium-dependent nuclear translocation in BHK cells, NF-AT4 is able to induce the transcriptional activation of an NF-AT reporter gene in these cells (Fig. 1g, h).

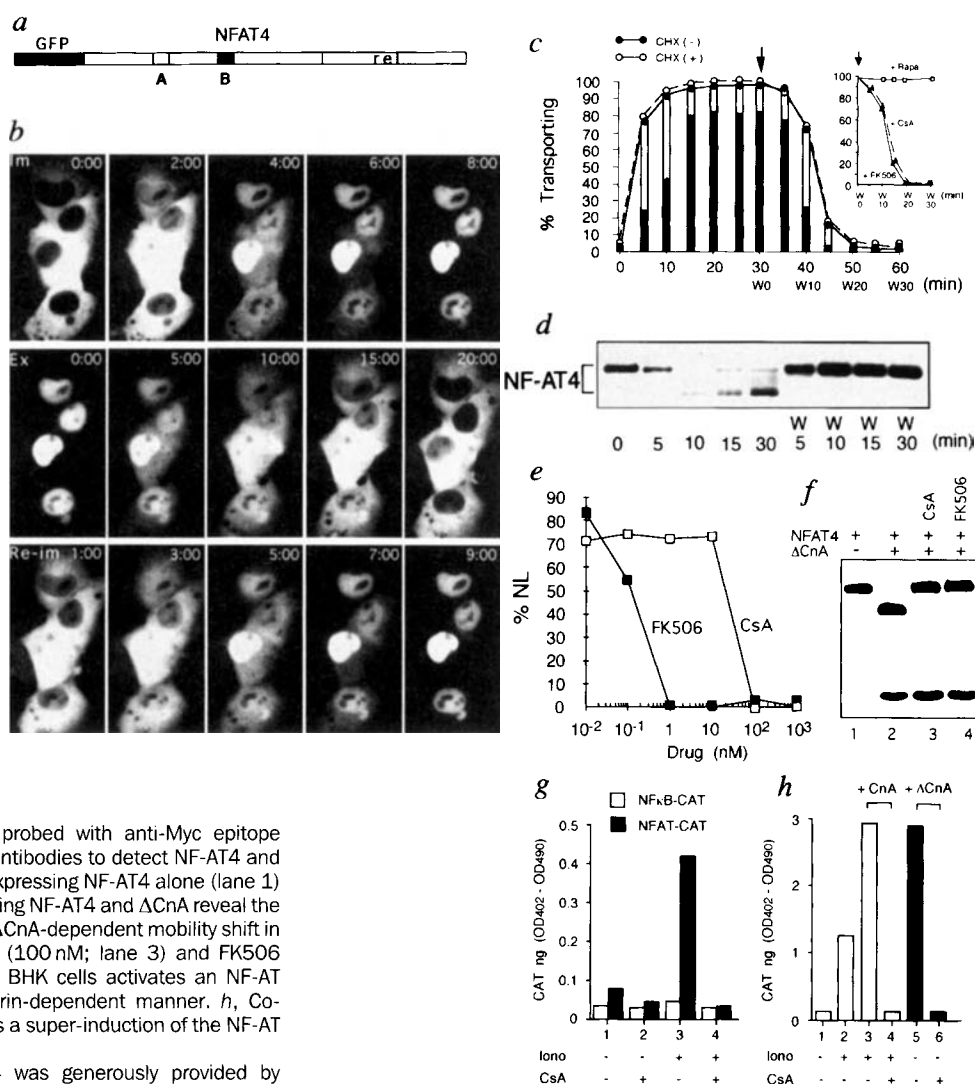
These results imply that only some of the dephosphorylation events on NF-AT4 are required for nuclear import, whereas others may have functional significance for NF-AT4 in the nucleus. A corollary of this finding is that calcineurin, previously thought to be a cytosolic phosphatase^{15,16}, may also function in the nucleus. To examine this possibility directly, we co-expressed calcineurin and NF-AT4 in U2OS cells. U2OS cells have low endogenous NF-AT4 translocation activity which is restored by calcineurin expression (Fig. 2a). Although calcineurin is restricted to the cytoplasm of unstimulated cells, it is localized to the cytoplasm and the nucleus upon calcium signalling (Fig. 2b). The kinetics of nuclear import of calcineurin were similar to those obtained for NF-AT4, with $t_{1/2} \sim 5$ –7 min (Fig. 2c). Nuclear export of calcineurin A (CnA) following ionophore removal gave a $t_{1/2}$ of 10–15 min (Fig. 2c). Δ CnA expressed alone appears to partition between the cytoplasm and the nucleus (Fig. 2d, a), whereas cells co-expressing Δ CnA and NF-AT4 showed an enhanced accumulation of Δ CnA in the nucleus (Fig. 2d, b and c). Remarkably, treatment of these cells with 100 nM CsA for 30 min results in a complete relocation of both Δ CnA and NF-AT4 from the nucleus to the cytoplasm (Fig. 2d, d and e). We also transfected U2OS cells with combinations of NF-AT4 and calcineurin cDNAs and analysed cell lysates for NF-AT4–CnA complexes by immunoprecipitating CnA. This analysis revealed a strong NF-AT4 co-precipitate under conditions where calcineurin was either activated by calcium ionophore (Fig. 3a, lane 3), or constitutively active, as with Δ CnA (Fig. 3a, lane 6). Significantly, no co-precipitation of NF-AT4 was detected from lysates of unstimulated cells expressing CnA (Fig. 3a, lane 2), or in lysates of cells treated with CsA or FK506 (Fig. 3a, lanes 4, 5, 7 and 8). These data confirm and extend observations of associations between calcineurin and a related NF-AT transcription factor, NF-AT1 (refs 11, 12). To localize the domain of NF-AT4 that interacts with calcineurin, we constructed two NF-AT4 mutants lacking either the Rel domain or the N terminus (Fig. 3b). When expressed in non-stimulated cells, these mutants and wild-type NF-AT4 runs as discrete species (Fig. 3c, lanes 1–3).

However, lysates of cells co-expressing Δ CnA and NF-AT4 mutants had a mobility shift of the NF-AT4 and the N-terminal domain (Fig. 3c, lanes 4, 6), suggesting that the N-terminal domain is the site of most of the dephosphorylation events on NF-AT4. Moreover, co-immunoprecipitation experiments showed strong interactions between calcineurin and either the full-length or the N terminus of NF-AT4, but not the C-terminal Rel domain (Fig. 3d).

FIG. 1 GFP–NF-AT4 nuclear import and export as a function of calcium signalling. **a**, Schematic of green fluorescent protein (GFP)–NF-AT4 fusion protein. **b**, Sequence of calcium-ionophore-dependent nuclear import (Im) and export (Ex) of GFP–NF-AT4 in transfected BHK cells. Bottom, re-import sequence. **c**, Kinetic analysis of NF-AT4 nuclear import and export in BHK cells. About 80% of these cells show complete nuclear localization (solid bars) of NF-AT4 at 15 min; the remaining 20% show a predominant, but not complete, nuclear staining (hatched bars). Identical analyses were done on transfected cells treated with 36 μ M cycloheximide (CHX) 30 min before and during calcium activation or washout (W) to block new protein synthesis (dotted line). Insert: a similar time course of nuclear export of NF-AT4 was obtained in medium containing either cyclosporin A (CsA; 100 nM) or FK506 (10 nM), while rapamycin (Rapa; 50 nM) did not promote export. **d**, Kinetic analysis of NF-AT4 mobility shifts during ionophore treatment and subsequent washout. **e**, BHK cells expressing NF-AT4 were incubated with titrating concentrations of FK506 and cyclosporin A and assayed for nuclear localization (NL) following 30 min of ionophore treatment. **f**, Immunoblot of BHK cells probed with anti-Myc epitope antibodies to detect NF-AT4 and anti-HA antibodies to detect Δ CnA. Cells expressing NF-AT4 alone (lane 1) show the slow-mobility form; those expressing NF-AT4 and Δ CnA reveal the fast-mobility form of NF-AT4 (lane 2). The Δ CnA-dependent mobility shift in NF-AT4 is blocked by both cyclosporin A (100 nM; lane 3) and FK506 (10 nM; lane 4). **g**, NF-AT4 expressed in BHK cells activates an NF-AT reporter gene in a calcium and calcineurin-dependent manner. **h**, Co-expressing of calcineurin and NF-AT4 yields a super-induction of the NF-AT reporter gene.

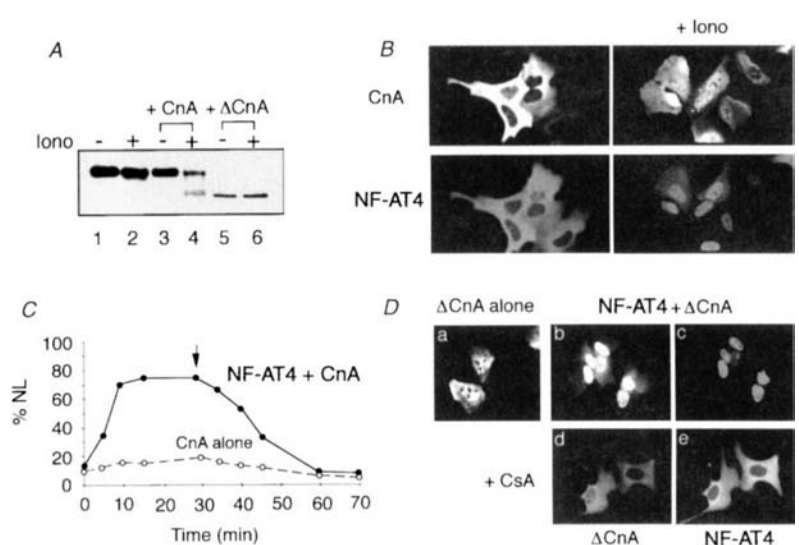
METHODS. The cDNA encoding NF-AT4 was generously provided by T. Hoey⁸. Vectors expressing a Myc-epitope-tagged NF-AT4 protein were constructed by PCR amplification of the NF-AT4 cDNA to introduce an *Xho*I site in 5' of codon 2 and an *Apa*I site after the natural stop codon, using oligonucleotides NF-AT4-1 (CCGCTCGAGACTACTGCAAACTGTGGC) and NF-AT4-2 (AAAGGGCCCTATAGTAAACCATCATG). After cleavage by *Xho*I and *Apa*I, the PCR product was cloned into the pCMV vector previously engineered to place a Myc epitope (EQKLISEEDL) at the N terminus²⁵. DNA transfections of baby-hamster kidney (BHK-21) and human osteosarcoma U2OS cells (American Type Culture Collection) were done using calcium phosphate precipitations essentially as described¹⁵. The coding sequence of the jellyfish *Aequorea victoria* green fluorescent protein (GFP)¹⁰ was mutated to yield amino-acid substitutions of T65 and A163 using standard techniques, and then amplified by PCR and inserted in-frame between the Myc-epitope tag and the N terminus of NF-AT4 in pCMV. BHK cells grown on 25-mm coverslips were transfected with the resulting vector (GFP–NF-AT4) and placed in a temperature-controlled chamber in DMEM (lacking phenol red) with 10% fetal calf serum and 25 mM HEPES, pH 7.4. Nuclear import of NF-AT4 was stimulated as described by adding 1 μ M A23187, and nuclear export of NF-AT4 followed after three rinses with fresh medium lacking A23187. GFP–NF-AT4 translocations were recorded using 50-ms exposures from a 100-W mercury source dampened to 5% with

To test the physiological significance of calcineurin–NF-AT4 interactions further, we produced and assayed a series of calcineurin active-site^{17,18} mutants (Δ CnA-H101Q, -H160Q and -H290Q) (Fig. 4a). Significantly, despite the loss of phosphatase activity, these mutants appear to retain their native structure, as judged by their ability to bind calcineurin B and to interact with immunophilin–drug complexes¹⁹ (data not shown). Unlike Δ CnA, all of the catalytically inactive calcineurin mutants are located



neutral density filters under FITC filters (HQ 450/50 excitation, Q 480 beam splitter, and HQ 510/50 emitter) on a Zeiss Axiovert microscope equipped with a Photometrics CCD camera containing a Kodak KAF1400 chip. The exposure sequences and imaging were controlled by MetaMorph (Universal Imaging) imaging software. Western blotting: for preparing whole cell lysates, coverslips were washed twice with ice-cold PBS, and cells extracted with 200 μ l preheated (70 °C) 2 $\times$ SDS-sample buffer containing 10% 2-mercaptoethanol. Lysates were transferred to microcentrifuge tubes, boiled for 5 min, and centrifuged at 10,000g for 10 min. 40 μ l of each sample was fractionated on 8% polyacrylamide–SDS gels and transferred to Immobilon membranes (Millipore), probed with antibodies, and developed using chemiluminescence reagents (ECL: Amersham). For transcriptional activation in BHK cells, the NF-AT-CAT reporter construct and NF- κ B-CAT reporter construct (kindly provided by L. Glimcher) were expressed together with NF-AT4 at a ratio $\sim$ 1:3. Chloramphenicol acetyltransferase (CAT) was assayed in cell lysates using an enzyme-coupled CAT antibody (Boehringer Mannheim).

FIG. 2 Nuclear localization of activated calcineurin. **A**, The NF-AT4 mobility shift defect in U2OS cells is rescued by calcineurin expression. NF-AT4 expressed in U2OS cells does not undergo a mobility shift in response to ionophore (lanes 1 and 2). In contrast, cells co-expressing NF-AT4 and either CnA²⁶ (lanes 3 and 4) or Δ CnA (lanes 5 and 6) show NF-AT4 mobility shifts that are ionophore-dependent and -independent, respectively. **B**, Immunofluorescence micrographs showing the cytoplasmic distribution of HA-tagged CnA and Myc-tagged NF-AT4 in unstimulated U2OS cells (left panels). After 30 min of ionophore treatment, the CnA signal is apparent in the nucleus, as is NF-AT4 in the same cells (right panels). **C**, Kinetic analysis of CnA nuclear import and export as a function of time of ionophore treatment. U2OS cells transfected with CnA alone (open circles) or CnA and NF-AT4 (closed circles) were stimulated with ionophore for the indicated times and then fixed and processed for immunofluorescent localization of CnA. 500 cells were scored for each time period for CnA nuclear import, with a positive end point determined by cells in which the nuclei displayed a signal equal to or greater than that of the cytoplasm. Nuclear export of CnA was assayed at various times after ionophore washout (arrow). **D**, Constitutively active Δ CnA is distributed to the cytoplasm and nucleus of unstimulated U2OS cells (a). In cells co-transfected with Δ CnA and NF-AT4, the extent of Δ CnA nuclear localization is constitutively higher than in cells expressing Δ CnA



alone (b). NF-AT4 appears exclusively nuclear in the same cells showing nuclear Δ CnA (c). Redistribution of both Δ CnA (d) and NF-AT4 (e) to the cytoplasm of U2OS cells following 30 min of treatment with cyclosporin A (100 nM).

exclusively in the cytoplasm, despite their co-expression with NF-AT4 (Fig. 4a, b). In addition, these mutants appear to act in a dose-dependent manner to suppress the ability of endogenous calcineurin to stimulate the translocation of NF-AT4 to the nucleus (Fig. 4a, b).

We show here that calcineurin functions in the unmasking of nuclear-localization signals (NLs) in NF-AT4, and is carried to the nucleus by virtue of its association with NF-AT4. In the nucleus, it is likely that this association with calcineurin is neces-

sary to maintain NF-AT4 in a transcriptionally active state for the duration of calcium signalling (Fig. 4c). An emerging concept in gene expression is the involvement of kinases in the regulation of transcription complexes. Many of the implicated kinases, including the ERKs, RSKs and JNKs, are themselves controlled at the level of nuclear import²⁰⁻²⁴. The possible involvement of a phosphatase in transcriptional control complements the notion that regulatory kinases impart direct control over transcription factors in the nucleus. Our results also highlight the activity of a powerful

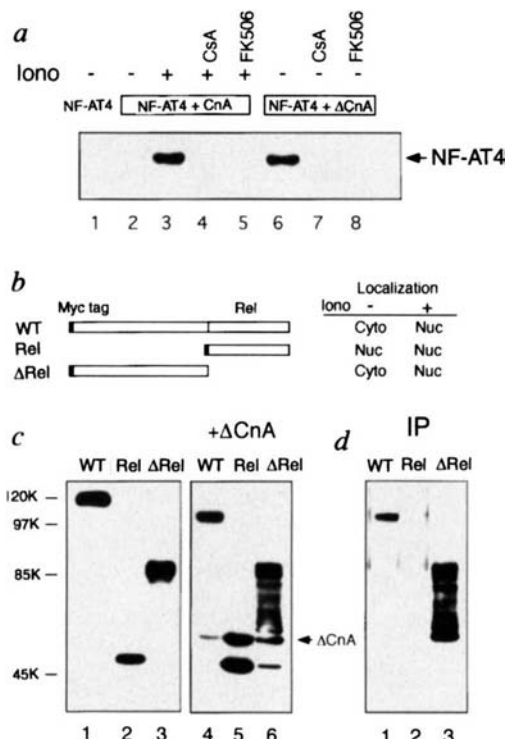


FIG. 3 a, Association of NF-AT4 with activated calcineurin. Lysates were prepared from U2OS cells expressing CnA alone (lane 1), NF-AT4 and CnA (lane 2), NF-AT4 and CnA after 30 min ionophore treatment (lane 3), NF-AT4 and CnA after 30 min cyclosporin A (lane 4), NF-AT4 and CnA after 30 min FK506 (lane 5), NF-AT4 and Δ CnA after 30 min cyclosporin A (lane 7), and NF-AT4 and Δ CnA after 30 min FK506 (lane 8), and soluble calcineurin immunoprecipitated with the anti-HA epitope antibody. **b**, Putative domain structure of NF-AT4 and subcellular localization of N- and C-terminal regions. The Rel domain lacking the N terminus is constitutively localized to the nucleus (Nuc), independent of ionophore treatment. The N terminus of NF-AT4 is cytoplasmic (Cyto) in unstimulated cells and moves to the nucleus upon ionophore treatment. WT, wild type. **c**, Whole-cell lysates of U2OS cells expressing NF-AT4 (lane 1), the NF-AT4 Rel domain (lane 2) or the NF-AT4 N terminus (lane 3) alone or with Δ CnA (lanes 4, 5 and 6) were probed with the anti-Myc and anti-HA epitope antibodies to detect NF-AT4 and Δ CnA, respectively. **d**, Detection of NF-AT4 in Δ CnA immunoprecipitates from cells co-expressing Δ CnA and NF-AT4 (lane 1), NF-AT4 Rel domain (lane 2), and the NF-AT4 N terminus (lane 3).

METHODS. NF-AT4 mutant constructs. The NF-AT4 mutant expression vectors, including the C- and N-terminal deletions, were constructed by PCR using oligonucleotides NF-AT4-1 and NF-AT4-3 (GATAGTGGGCCCTATGAACACAGCTCTAATTTTCC), and NF-AT4-2 and NF-AT4-4 (CCGCTCGAGGATGACCAAGGGAGTTTA), respectively. The C-terminal deletion mutant is a protein with a Myc-epitope followed by amino acids 2 to 359 of NF-AT4 (ref. 8). The N-terminal deletion mutant encodes a protein with a Myc-epitope followed by the Rel domain (amino acids 359 to 708).

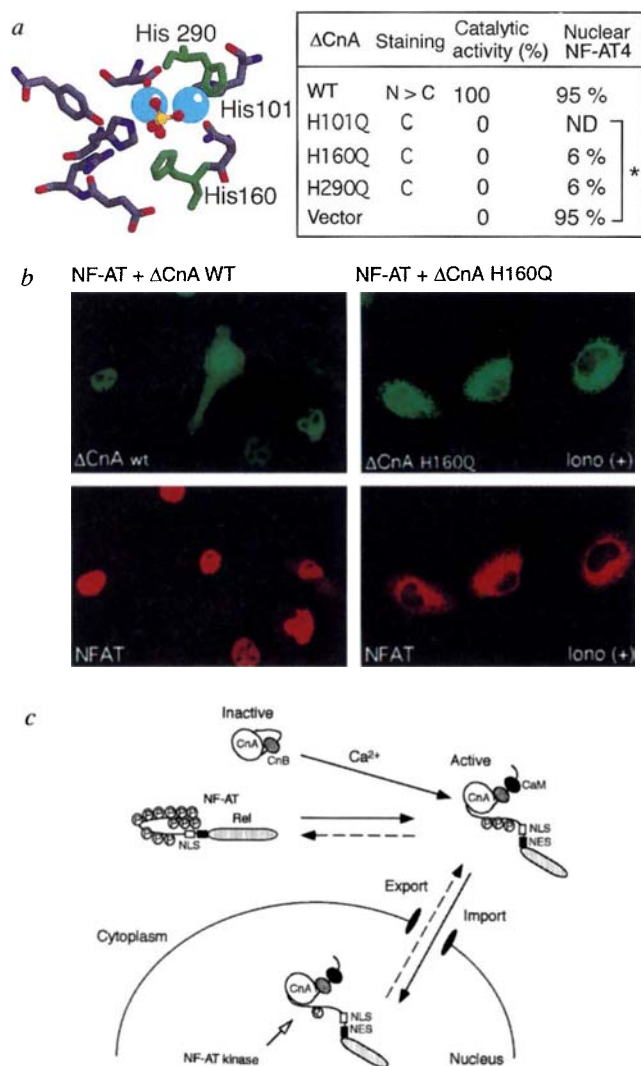


FIG. 4 Catalytically inactive calcineurin mutants are exclusively cytoplasmic and interfere with NF-AT4 translocation. **a**, Model of calcineurin active site based on conserved residues in protein phosphatase (ref. 17). The three histidine residues (at positions 101, 160 and 290) mutated to glutamine are labelled. The two metal ions are shown in blue, adjacent to a phosphate group. These mutants were synthesized *in vitro* and tested for catalytic activity towards the RII substrate¹⁹. The three inactive ΔCnA mutants are located exclusively in the cytoplasm of transfected cells, as judged by immunofluorescence. When co-transfected with NF-AT4, the H160Q and H290Q mutants suppress the ability of endogenous calcineurin to translocate NF-AT4 when stimulated with ionomycin (asterisk). **b**, Wild-type ΔCnA is primarily nuclear and induces the nuclear translocation of NF-AT4 (left panels). In contrast, ΔCnA-H160Q is exclusively cytoplasmic and fails to translocate NF-AT4 (right panel). **c**, Calcium signalling through calcineurin and NF-AT. Initially, both calcineurin (CnA and CnB), shown in its auto-inhibited form, and NF-AT4, with masked nuclear location signals, reside in the cytoplasm. Calcium activates calcineurin, which then binds to and dephosphorylates NF-AT4, causing an unmasking of its nuclear location signals. Both NF-AT4, and the associated CnA are translocated to the nucleus, where they encounter a putative NF-AT4 kinase. As calcineurin activity decreases at the end of calcium signalling, the NF-AT4 kinase predominates and promotes the nuclear export of NF-AT4. NLS and NES refers to putative nuclear localization and export signals, respectively.

METHODS. Mutagenesis of ΔCnA and analysis of the catalytic activity of mutants were done essentially as described¹⁹. For analysis of dominant-negative activity of the calcineurin mutants, BHK cells were transfected with either NF-AT4 alone or with plasmid expressing mutant ΔCnA and CnB. After 16 h expression, cells were treated with ionophore for 30 min and processed for immunofluorescence to assay for NF-AT4 translocation.

NF-AT4 kinase involved in the attenuation of calcium signalling in the nucleus. □

Received 13 May; accepted 14 June 1996.

- Schreiber, S. L. & Crabtree, G. R. *Immun. Today* **13**, 136–142 (1992).
- Crabtree, G. R. & Clipstone, N. A. *Rev. Biochem. Chem.* **63**, 1045–1083 (1994).
- Rao, A. *Immun. Today* **15**, 274–281 (1994).
- Flanagan, W. M., Corthesy, B., Bram, R. J. & Crabtree, G. R. *Nature* **352**, 803–807 (1991).
- Liu, J. et al. *Cell* **66**, 807–815 (1991).
- Ruff, V. A. & Leach, K. L. *J. Biol. Chem.* **270**, 22602–22607 (1995).
- Shaw, K. T.-Y. et al. *Proc. natn. Acad. Sci. U.S.A.* **92**, 11205–11209 (1995).
- Hoey, T., Sun, Y. L., Williamson, K. & Xu, X. *Immunity* **2**, 461–472 (1995).
- Masuda, E. S. et al. *Molec. cell. Biol.* **15**, 2697–2706 (1995).
- Chalfie, M., Tu, Y., Euskirchen, G., Ward, D. D. & Prasher, D. C. *Science* **263**, 802–805 (1994).
- Wesselborg, S., Fruman, D. A., Sagoo, J. K., Bierer, B. E. & Burakoff, S. J. *J. Biol. Chem.* **271**, 1274–1277 (1996).
- Loh, C. et al. *J. Biol. Chem.* **271**, 10884–10891 (1996).
- Hubbard, M. J. & Klee, C. B. *Biochemistry* **28**, 1868–1874 (1989).
- O'Keefe, S. J., Tamura, J., Kincaid, R. L., Tocci, M. J. & O'Neill, E. A. *Nature* **357**, 692–694 (1992).
- Shibasaki, F. & McKeon, F. *J. Cell Biol.* **131**, 735–743 (1995).
- Klee, C. B., Draetta, G. F. & Hubbard, M. J. *Adv. Enzym. relat. Areas molec. Biol.* **61**, 149–200 (1988).
- Goldberg, J. et al. *Nature* **376**, 745–753 (1995).
- Griffith, J. P. et al. *Cell* **82**, 507–522 (1995).
- Milan, D., Griffith, J., Su, M., Price, E. R. & McKeon, F. *Cell* **79**, 437–447 (1994).
- Karin, M. & Hunter, T. *Curr. Biol.* **5**, 747–757 (1995).
- Chen, R. H., Sarnacki, C. & Blenis, J. *Molec. cell. Biol.* **12**, 915–927 (1992).
- Tyers, M. & Fletcher, B. *Molec. cell. Biol.* **13**, 5659–5669 (1993).
- Peter, M., Garter, A., Horecka, J., Ammerer, G. & Herskowitz, I. *Cell* **73**, 747–760 (1993).
- Hibi, M., Lin, A., Smeal, T., Minden, A. & Karin, M. *Genes Dev.* **7**, 2135–2148 (1993).
- Heald, R., McLoughlin, M. & McKeon, F. *Cell* **74**, 463–474 (1993).
- Guerni, D. & Klee, C. B. *Proc. natn. Acad. Sci. U.S.A.* **86**, 9183–9187 (1989).

ACKNOWLEDGEMENTS. We thank T. Hoey and S. McKnight for NF-AT4 cDNA; J. Rooney, M. Hodge and L. Glimcher for advice on NF-AT transactivation; M. Chalfie, T. Bestor, P. Silver and L. Lanini for reagents and advice on construction of GFP–NF-AT4; L. Ma, R. Davis and M. Kirschner for instruction and use of their imaging systems; J. Goldberg for protein phosphatase-1 coordinates; M. Shibata, S. Kitamura and associates at MBL for antibodies; B. Bierer, A. Rao, A. Yang, C. Walsh, B. Sleckman and H. Green for discussion; and all members of our laboratory for reagents and advice. This work was supported by a fellowship to F.S. from the Human Frontier Science Program and grants from the NIH and the American Cancer Society to F.M.

CORRESPONDENCE and requests for materials should be addressed to (e-mail: mckeon@warren.med.harvard.edu).

RNA-catalysed RNA polymerization using nucleoside triphosphates

Eric H. Eklund & David P. Bartel

Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142, USA

THE hypothesis that certain RNA molecules may be able to catalyse RNA replication is central to current theories of the early evolution of life^{1–6}. In support of this idea, we describe here an RNA that synthesizes RNA using the same reaction as that employed by protein enzymes that catalyse RNA polymerization. In the presence of the appropriate template RNA and nucleoside triphosphates, the ribozyme extends an RNA primer by successive addition of up to six mononucleotides. The added nucleotides are joined to the growing RNA chain by 3',5'-phosphodiester linkages. The ribozyme shows marked template fidelity: extension by nucleotides complementary to the template is up to 1,000 times more efficient than is extension by mismatched nucleotides.

Ribozymes that ligate RNA have been found in nature^{7–11} and have been isolated from large pools of random RNA sequences^{12,13}. Of particular interest for exploring the possibility of self-replicating RNA are ligase ribozymes that, like RNA polymerases, join RNA substrates using reactions in which pyrophosphate (PP_i) is displaced^{11,12}. The class I ligase, a ribozyme derived from random RNA sequences, efficiently catalyses a

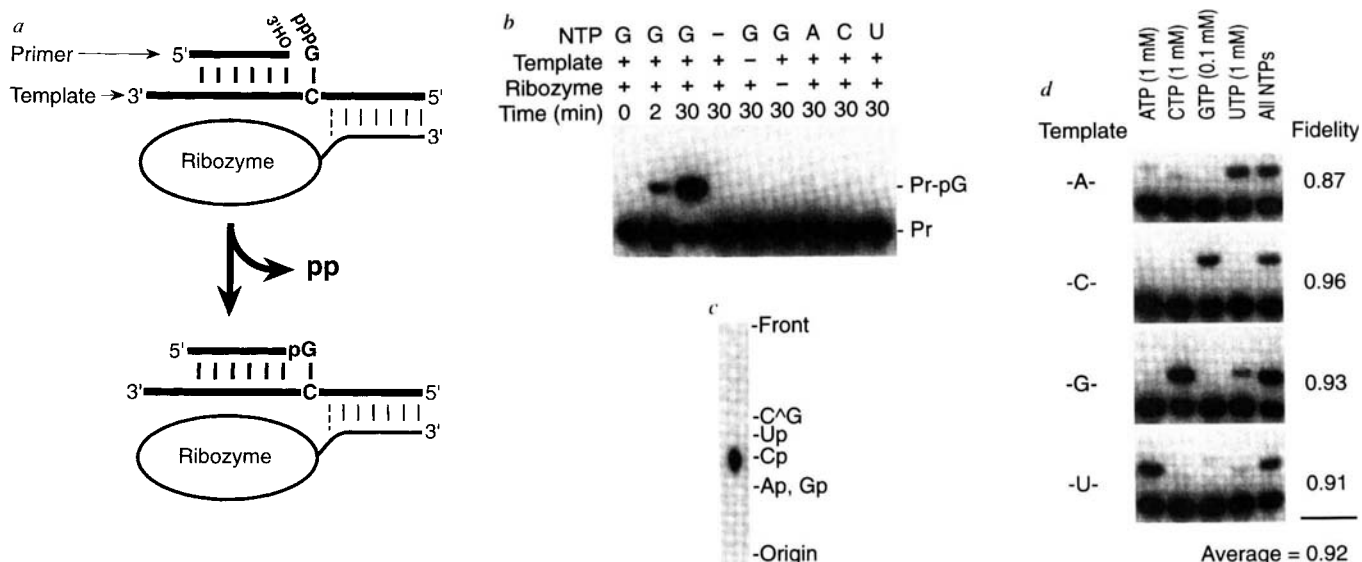


FIG. 1 Single nucleotide addition. **a**, Design of RNAs for addition of GMP using a -C- template. The ribozyme binds the template-primer, in part through base pairing of the 5' end of the template with the 3' tail of the ribozyme (thin vertical lines; dashed line denotes a proposed A-A pair¹⁵). GTP (pppG) binds the primer-template-ribozyme ternary complex such that the primer and the GTP are aligned by contiguous Watson-Crick pairing (thick vertical lines) to the template RNA. The ribozyme promotes attack of the primer 3' hydroxyl (3'HO) on the GTP α -phosphate, resulting in the extension of the primer by one nucleotide (pG), with concomitant release of inorganic pyrophosphate (PP). **b**, Time course of GMP addition. ³²P-labelled primer (Pr) was incubated with or without the indicated NTP (1 mM), template and ribozyme. Extended primer (Pr-pG) was separated from non-extended primer on a sequencing gel (20% acrylamide/8 M urea). **c**, Nuclease digestion of the extended primer indicates that the added nucleotide is joined to the primer by a 3',5'-phosphodiester linkage. Unlabelled primer extended using [³²P]GTP was gel-purified, mixed with carrier RNAs, and digested with nuclease T2, which does not efficiently digest 2',5' RNA linkages. Digestion products were separated by thin-layer chromatography on PEI cellulose, developed with 1 M LiCl. Ultraviolet shadowing revealed digestion products of carrier tRNA (nucleoside 3'-monophosphates: Ap, Cp, Gp, Up) as well as a 2',5'-linked carrier dinucleotide (C^G). **d**, Observed Watson-Crick fidelity of primer extension with an optimized ratio of NTP concentrations. Gel slices show the amount of extension after 30 min (-C- and -G- templates) or 180 min (-A- and -U- templates) with individual nucleoside triphosphates and with the combina-

tion of four NTPs. Extension rates were determined using these time points, as well as additional time points not shown. The observed fidelity was calculated as the extension rate for the Watson-Crick match divided by the sum of the extension rates for all four individual NTP possibilities. Note that this observed fidelity differs from the standard fidelity (Table 1) in that the observed fidelity is calculated from the extension rates when using a specific ratio of the four NTPs, whereas the standard fidelity is derived from the k_{cat}/K_m^{NTP} values.

METHODS. Reactions were done at 22 °C in 30 mM Tris-HCl, pH 8.0, 200 mM KCl, 60 mM MgCl₂ and 0.6 mM EDTA, with 2.5 μ M ribozyme, 2.0 μ M template, 1 μ M primer and the indicated concentration of NTP, using described methods¹⁵. Ribozyme construct b1-233t was generated as described¹⁵ (b1-233t sequence: 5'-ggaaaagac aaacugccc ucagagcuug agaacaucuu cggaugcaga ggaggcagcc uccguggcu uuaacgcca cguucuaac aaugcca, where underline represents segment pairing with template RNA). *In vitro* transcription³⁰ was used to generate template RNAs (5'-ggcuauaNgacuggaacca; single underline, segment pairing with ribozyme; double underline, segment pairing with primer; N, template base designed to pair with the nucleoside triphosphate, that is, for the -A- template, N = A). A synthetic DNA-RNA chimera served as the primer (5'-AAAccaguc; where upper-case represents DNA, and double underline, segment pairing with template)¹⁵.

reaction in which the 3'-hydroxyl of one substrate RNA attacks the 5'-triphosphate of another substrate molecule, forming a new phosphodiester linkage with concomitant release of pyrophosphate^{14,15}. To explore this ribozyme's potential as a starting point for developing an RNA polymerase, we examined whether it could utilize a mononucleoside triphosphate rather than an oligonucleotide triphosphate as one of its substrates (Fig. 1a). When incubated with a primer RNA, GTP, and a template RNA designed to align the primer with the GTP by means of Watson-Crick pairing, the ribozyme elongated the primer by addition of GMP (Fig. 1b). When assayed over a range of GTP concentrations, the reaction showed saturable kinetics that were consistent with a single GTP-binding site (Michaelis-Menten parameters, $K_m^{GTP} = 5$ mM, $k_{cat} = 0.3$ min⁻¹). In the presence of excess primer-template, catalytic turnover was observed (7.5 turnovers in one hour, using 0.1 μ M ribozyme, 5 μ M primer, 5 μ M template, 3 mM GTP). Analysis of the product confirmed that GMP was joined to the primer by a 3',5'-phosphodiester linkage (Fig. 1c).

The primer extension reaction was highly specific for elongation using GTP, the Watson-Crick match to the template. The other three nucleotides added 600 to 1,000 times less efficiently (Fig. 1b). When the template RNA was substituted with templates designed to code for other nucleotides, addition by the Watson-

Crick match was always favoured over addition by mismatch possibilities (Table 1). The k_{cat}/K_m^{NTP} values for all 16 template-NTP (where NTP is any nucleoside triphosphate) combinations can be used to calculate a standard Watson-Crick fidelity of 0.85 (Table 1). In other words, if given all four NTPs at equimolar concentrations, extension by the matched nucleotide typically would comprise 85% of the total extension. Certain asymmetric NTP ratios produced observed fidelities significantly greater than 0.85; the observed fidelity averaged 0.92 when the GTP concentration was uniformly lowered to one tenth of the concentration of the other three NTPs (Fig. 1d). At an observed fidelity of 0.92, extension by the Watson-Crick combination is, on average, 34 times faster than that by each of the three non-Watson-Crick combinations (0.92, compared to 0.027).

To examine the ribozyme's ability to extend a primer by more than one nucleotide, two additional template bases were inserted in the region between the primer-template helix and the template-ribozyme helix (Fig. 2a). This insertion makes three template residues available to specify the successive addition of up to three nucleotides. The template was also joined to the 5' terminus of the ribozyme; tethering of the template to the ribozyme typically enhances extension efficiency by 2-6-fold. When incubated with all four NTPs and primer, such ribozyme-template

fusions catalysed primer extension by three nucleotides (Fig. 2). As polymerization proceeds, the primer-template helix shifts relative to the site of chemistry, changing the identity of the Watson-Crick pairs at most positions in the helix. With most templates shown in Fig. 2, none of the primer-template base pairs (counting back from the 3' terminus of the growing strand) is the same during all three elongation steps. This shows that, as with polymerases found in nature, there is no absolute requirement for any particular Watson-Crick pair at any position of the primer-template helix.

The accuracy of extension was reassessed using the ribozyme-template format that permits multiple nucleotide addition. Although extension slowed considerably when switching from the single-nucleotide-addition format, the average fidelity of polymerization across from the four homopolymeric template segments (0.88; Fig. 2e) was similar to that observed for single nucleotide addition (0.92; Fig. 1d). Another similarity between the two formats was that correct extension directed by G or C template residues was 10 to 40 times more efficient than that directed by A or U residues (Table 1 and Fig. 2).

Although much lower than the ≥ 0.996 fidelity seen with viral polymerases that synthesize RNA using RNA templates^{16,17}, the observed fidelity is significantly greater than would be expected from the intrinsic stability of Watson-Crick matches compared with the stability of mismatches. Watson-Crick pairs at the ends of RNA helices are sometimes no more stable than are mismatches¹⁸. Using the relative stabilities of terminal base pairs and mismatches¹⁸ to predict the fidelity of polymerization yields a maximal (using optimal NTP ratios) fidelity averaging about 0.40—a mere twofold advantage of Watson-Crick matches over non-Watson-Crick combinations (0.40, compared to an average of 0.20 for each of the three mismatches).

Use of nucleoside triphosphates to extend an RNA or DNA molecule is a fundamental reaction in all living organisms. The reaction shown in Fig. 2 is similar to that of a telomerase, a specialized DNA polymerase that uses a short segment of a built-in RNA template to direct the extension of chromosome ends. It has been proposed that before the advent of protein catalysis, a ribozyme with activity similar to that of the transfer RNA nucleotidyltransferase served as a telomerase¹⁹. Although tRNA nucleotidyltransferase is primarily responsible for adding and maintaining the CCA-3' tail of the tRNA acceptor stem, this cellular enzyme can function as a simple telomerase for viral genomic RNAs that assume tRNA-like structures at their 3' ends^{20,21}. To the extent that our ribozyme can add a CCA-3' tail to an RNA primer (Fig. 2c), we confirm that RNA can catalyse this reaction which is thought to have been important during the RNA world.

More central to the RNA-world hypothesis than an RNA telomerase is an RNA replicase, an RNA polymerase ribozyme which, with the aid of an RNA template encoding a second polymerase molecule, would be capable of autocatalytic replication¹⁻⁶. Previous efforts to design ribozymes with polymerase properties have focused on derivatives of self-splicing introns that promote RNA oligonucleotide disproportionation reactions, where one RNA molecule becomes longer at the expense of a second RNA. Although progress has been made by this approach²²⁻²⁷, reactions using self-splicing intron derivatives still suffer from low Watson-Crick fidelity; the highest fidelity achieved thus far has been 0.65 (ref. 28). More important, polymerization by oligonucleotide disproportionation is intrinsically prone to undesirable side reactions, which has led to the assertion that the substrates of a replicase ribozyme must have been more analogous to those of contemporary polymerases; that is, they must have been activated by a moiety like pyrophosphate rather than by RNA⁶.

The RNA polymerization in Fig. 2 demonstrates features essential for polymerization by an RNA replicase, such as template dependence, use of appropriately activated mononucleotides, regioselectivity for the formation of 3',5'-phosphodiester

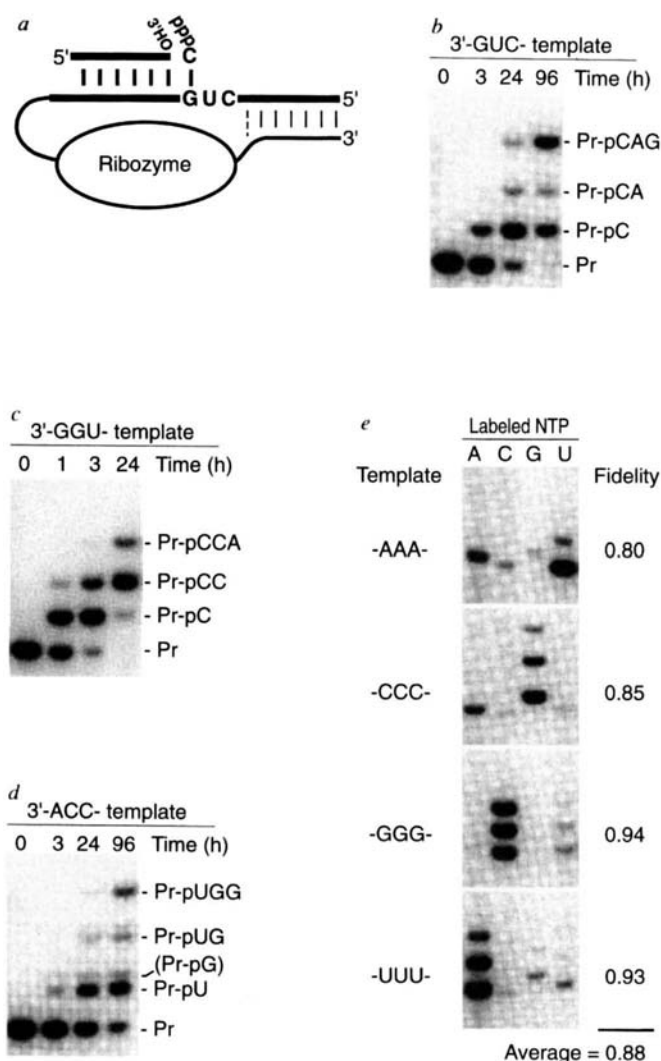


FIG. 2 RNA polymerization. **a**, Design of a ribozyme-template fusion for sequential addition of three nucleotides encoded by a 3'-GUC- template segment. **b-d**, Time course of polymerization using 32 P-labelled primer, a mixture of all four NTPs, and the indicated template segment. The identities of the indicated extension products were confirmed by comparison to reactions in which individual nucleotides were omitted (data not shown). In some cases, minor bands can be detected that represent incorporation by a non-Watson-Crick alternative (for example, band Pr-pG in **d**). **e**, Observed Watson-Crick fidelity of polymerization when using the four homopolymeric template segments. Reactions were as in **b-d**, except that the primer was not labelled and the NTP mixture was spiked with the indicated [32 P]NTP (final specific activity, 1.6 Ci mmol⁻¹). Gel slices show the amount of extension after 3 h (-CCC- and -GGG- templates) and 24 h (-AAA- and -UUU- templates).

METHODS. Reactions were done as for Fig. 1b, except that they contained all four NTPs (1 mM GTP; 10 mM ATP, CTP, UTP) and an additional 30 mM MgCl₂ to compensate for the Mg²⁺ expected to be complexed by the added NTPs. Ribozyme constructs were the same as construct b1-233t (Fig. 1), except the 5' terminal residues (G1 and G2) were replaced with a template segment (5'-ggcuauaNNNgacugguacaa, where the single underline represents segment pairing with the ribozyme 3' terminus, and double underline, segment pairing with primer; NNN are template bases designed to specify the nucleoside triphosphates that extend the primer). After stopping the reactions, samples were incubated (at 80 °C for 3 min) with a competitor DNA (5'-attgtctttttgtaccagtcNNNtatagcca) designed to hybridize with the template RNA. Incubation with competitor DNA led to better gel resolution because it prevented reassociation of the extended primer with the template. For each template in **e**, the observed fidelity was calculated as the incorporation by the nucleotide matching the template, divided by the incorporation by all four nucleotides.

TABLE 1 Template dependence of single nucleotide addition

Template	Relative efficiencies of extension*				Standard fidelity
	ATP	CTP	GTP	UTP	
-A-	0.062	0.054	0.12	1.0 (1.80)	0.81
-C-	0.0010	0.0014	1.0 (68)	0.0013	0.996
-G-	0.0017	1.0 (28)	0.0017	0.071	0.93
-U-	1.0 (2.5)	0.015	0.34	0.048	0.71
				Average	0.85

Reactions were performed as for Fig. 1. For each template–NTP combination, the apparent second-order rate constant, k_{cat}/K_m^{NTP} , was determined from the slope of the line plotting observed rate constants as a function of NTP concentration (using NTP concentrations at least 10 times lower than the K_m^{NTP}). For each template, the four k_{cat}/K_m^{NTP} values were normalized to the k_{cat}/K_m^{NTP} value of the Watson–Crick match, yielding the relative efficiencies of addition. (The relative efficiency of addition for a mismatch is the same as its error frequency²⁹.) Standard fidelity was calculated as the k_{cat}/K_m^{NTP} value of the Watson–Crick match, divided by the sum of the k_{cat}/K_m^{NTP} values for all four NTP possibilities. The average fidelity is the geometric average of the fidelities. (Because the overall accuracy of RNA polymerization is represented by the product of the fidelities for each added nucleotide, geometric averages are reported throughout this study.)

* The k_{cat}/K_m^{NTP} values ($M^{-1} min^{-1}$) for Watson–Crick matches are shown in parentheses.

bonds, and absence of sequence requirements within the primer–template helix. However, it is deficient in two key respects: the template is covalently linked to the catalyst, and polymerization beyond three nucleotides is blocked, presumably by specific base pairing of the template to the ribozyme. The single-nucleotide-addition experiments demonstrate that the ribozyme can function with an external rather than a tethered template (Fig. 1). The block to polymerization by the ribozyme–template helix is more difficult to overcome, but it was partially circumvented when

template and ribozyme were designed to base-pair in multiple registers (Fig. 3a). When using this strategy with an external template, polymerization of up to six nucleotides was detected after a long (6-day) incubation (Fig. 3b).

The progress made in deriving an RNA polymerase from random RNA sequences bodes well for the eventual demonstration of autocatalytic RNA replication. It will be of interest to examine the extent to which *in vitro* evolution and engineering efforts can improve the properties of our ribozyme. Extended copying of any RNA sequence may be possible if the 5' portion of the template was recognized primarily by sugar–phosphate contacts rather than by pairing to template bases. Other critical improvements would include more avid NTP binding and increased template fidelity. □

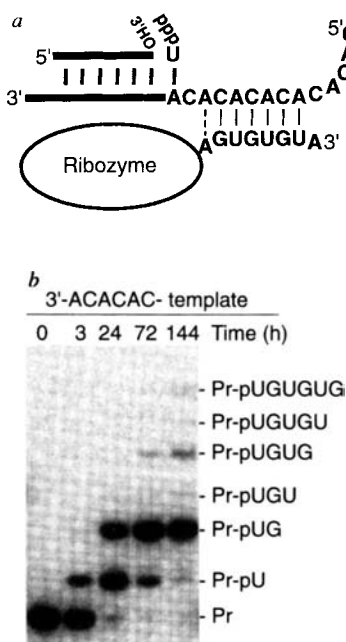


FIG. 3 RNA polymerization of up to 6 nucleotides. *a*, Design of template and ribozyme sequences such that the external template can pair with the ribozyme in multiple registers. *b*, Time course of polymerization. The reaction was done as for Fig. 1b, but contained UTP (20 mM), GTP (2 mM) and an additional 60 mM $MgCl_2$. The ribozyme was the same as construct b1-233t (Fig. 1), except that the 3' terminal segment (5'-aaugcca) was replaced by (5'-agugugua). The template RNA was 5'-gacacacacacagacuggaacca. After stopping the reactions, samples were heated in the presence of the appropriate competitor DNA, as for Fig. 2.

Received 2 February; accepted 16 May 1996.

- Sharp, P. A. *Cell* **42**, 397–400 (1985).
- Pace, N. R. & Marsh, T. L. *Orig. Life* **16**, 97–116 (1985).
- Cech, T. R. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4360–4363 (1986).
- Orgel, L. E. *J. theor. Biol.* **123**, 127–149 (1986).
- Joyce, G. F. *Nature* **338**, 217–224 (1989).
- Joyce, G. F. & Orgel, L. E. in *The RNA World* (eds Gestland, R. & Atkins, J.) 1–25 (Cold Spring Harbor Press, Cold Spring Harbor, New York, 1993).
- Kruger, K. et al. *Cell* **31**, 147–157 (1982).
- Peebles, C. L. et al. *Cell* **44**, 213–223 (1986).
- Van Tol, H., Buzayan, J. M. & Bruening, G. *Virology* **180**, 23–30 (1991).
- Saville, B. J. & Collins, R. A. *Proc. natn. Acad. Sci. U.S.A.* **88**, 8826–8830 (1991).
- Mori, M., Niemer, I. & Schmelzer, C. *Cell* **70**, 803–810 (1992).
- Bartel, D. P. & Szostak, J. W. *Science* **261**, 1411–1418 (1993).
- Chapman, K. B. & Szostak, J. W. *Chem. Biol.* **2**, 325–333 (1995).
- Eklund, E. H., Szostak, J. W. & Bartel, D. P. *Science* **269**, 364–370 (1995).
- Eklund, E. H. & Bartel, D. P. *Nucleic Acids Res.* **23**, 3231–3238 (1995).
- Steinhauer, D. A. & Holland, J. J. *J. Virol.* **57**, 219–228 (1986).
- Ward, C. D. & Flanagan, J. B. *J. Virol.* **66**, 3784–3793 (1992).
- Turner, D. H., Sugimoto, N. & Freier, S. M. *Rev. Biophys. biophys. Chem.* **17**, 167–192 (1988).
- Maizels, N. & Weiner, A. M. *Proc. natn. Acad. Sci. U.S.A.* **91**, 6729–6734 (1994).
- Joshi, R. L., Joshi, S., Chapelle, F. & Haenni, A. L. *EMBO J.* **2**, 1123–1127 (1983).
- Rao, A. L., Dreher, T. W., Marsh, L. E. & Hall, T. C. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5335–5339 (1989).
- Been, M. D. & Cech, T. R. *Science* **239**, 1412–1416 (1988).
- Doudna, J. A. & Szostak, J. W. *Nature* **339**, 519–522 (1989).
- Doudna, J. A., Couture, S. & Szostak, J. W. *Science* **251**, 1605–1608 (1991).
- Green, R. & Szostak, J. W. *Science* **258**, 1910–1915 (1992).
- Chowrira, B. M., Berzal-Herranz, A. & Burke, J. M. *EMBO J.* **12**, 3599–3605 (1993).
- Mueller, M. W., Hetzer, M. & Schweyen, R. J. *Science* **261**, 1035–1038 (1993).
- Bartel, D. P., Doudna, J. A., Usman, N. & Szostak, J. W. *Molec. cell. Biol.* **11**, 3390–3394 (1991).
- Wong, I., Patel, S. S. & Johnson, K. A. *Biochemistry* **30**, 526–537 (1991).
- Milligan, J. F. & Uhlenbeck, O. C. *Meth. Enzym.* **180**, 51–62 (1989).

ACKNOWLEDGEMENTS. We thank K. Williams, M. Moore, T. Tuschl and J. Theriot for helpful comments on the manuscript. This work was supported by a grant from the NIH and by a grant from the W. M. Keck foundation to the Whitehead Fellows program.

CORRESPONDENCE and requests for materials should be addressed to D.P.B.

Labelling techniques

This issue covers the reagents and equipment necessary for cellular and molecular imaging and detection, including a luminometer, a phosphor imaging system, reagents to detect intracellular cytokines and alternatives to radiolabelling.

VAYTEK has developed CoLocalization for those who use multiple fluorescent probes in capturing a separate image or images for each probe colour. This new software module should allow researchers to determine the amount of overlap between the images. Quantitative measurements are said to be easy to obtain with this module. For example, the percentage of overlap and the probability that a pattern is greater than chance can be determined. The system requires a Pentium computer running either Windows 3.1 or Windows95 with at least 16 MB of RAM. VayTek's modular approach should make it easy to retrofit, customize and upgrade image analysis systems.

Reader Enquiry No. 100

Reagents and kits

A new non-radioactive ELISA that is designed to provide an accurate measurement of true cellular proliferation is now available from Amersham. Using a highly specific anti-5-bromo-2'-deoxyuridine (BrdU) conjugate, the Biotrak cell proliferation ELISA offers stated detection levels of 100 cells per well in a microtitre-plate format. The ELISA should be useful



Biotrak cell proliferation ELISA.

for studying a variety of areas, including the effects of growth factors and cytokines; the determination of inhibitory or stimulatory effects of compounds in environmental, biomedical and pharmaceutical research; and the determination of the chemosensitivity of tumour cells to different cytostatic drugs. Depending on the cell culture, the ELISA can be completed in just over two hours. Supplied as a fully functional test kit, the ELISA includes sufficient material to measure ten 96-well plates.

Reader Enquiry No. 101

R&D Systems now offers reagents to detect intracellular cytokines. These monoclonal antibodies are now available for the detection of intracytoplasmic forms of cytokine by flow cytometry, conjugated to phycoerythrin or fluorescein where necessary. Once the cells of interest have been fixed and permeabilized, staining reveals how cellular activation can result in cytokine synthesis. The reagents are particularly useful for studying cytokine production by T-cell subsets, thereby allowing determination of whether a type-1 or type-2 response has been elicited. Antibodies are available for the detection of intracellular IL-2, IL-4, IL6, TNF- α and IFN- γ . In addition, conjugated negative controls are available for the verification of results. The ability to investigate cytokine expression at the single-cell level, by combining multiparameter analysis with the speed of flow cytometry, is likely to offer new insights into immunity.

Reader Enquiry No. 102

Boehringer Mannheim has tackled the problems associated with traditional radioactive ligand labelling methods (safety, disposal and expensive reagents) with the launch of the receptor binding analysis kit. The kit utilizes a high-efficiency labelling technique to isolate ligand-receptor complexes. It uses a new series of tetrafluorophenyl azide reagents (Atf) to eliminate the safety and environmental concerns posed by radiochemical methods and reduce the amount of time spent on training, clean-up and the administration associated with radioisotope usage. The kit is designed for convenience, with ready-to-use reagents for one-step ligand-receptor binding. The complete kit features a labelling reagent for ligands containing the Atf photoreactive residue, an active ester and a biotin label. The trifunctional labelling reagent reacts with the ligand and results in a photoactivated ligand. When activated by ultraviolet light, the photoactivated ligand is then covalently linked to the binding region of its specific receptor. The kit also contains immobilized streptavidin, which acts as a purification matrix for the isolation of the photoaffinity biotin-labelled, ligand-receptor complexes. Applications for the receptor binding analysis kit include the study of ligand-receptor, antigen-antibody and interleukin interactions.

Reader Enquiry No. 103

Sigma Immunochemicals has introduced secondary antibodies and avidin reagents conjugated to Cy3 for use in immunohistochemistry and immunocytochemistry. The available antibodies labelled with Cy3 are goat anti-human IgG, goat anti-human IgM, rabbit anti-goat IgG, sheep anti-mouse IgG and sheep anti-rabbit IgG. The avidin reagents available are ExtrAvidin and streptavidin.

Cy3 is a hydrophilic, red fluorescent dye and represents a new generation of fluorophores that are said to cause less nonspecific binding and aggregation. It is more photostable and quenching is minimized even at high F/P ratios. Maximally excited at 554 nm, Cy3 emits at 568–574 nm and can be used with traditional TRITC filter arrangements. The Cy3-labelled avidin reagents should prove useful for detecting biotin-labelled probes by fluorescence *in situ* hybridization. Monoclonal anti-smooth muscle actin (clone 1A4) is also available labelled with Cy3.

Reader Enquiry No. 104

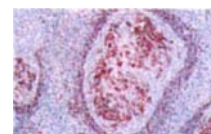
The Dako CSA system HRP is a sensitive immunohistochemical staining procedure for monoclonal antibodies, which utilizes an enzyme-catalysed deposition of biotin, resulting in an amplified signal that allows the detection of small quantities of reactive antigen. The system enables certain monoclonal antibodies, which are normally not compatible with formalin-fixed, paraffin-embedded sections, to be used on such specimens. The company is also developing a line of prediluted antibodies for use with the CSA system, which includes the option of a ready-to-use antibody.

Reader Enquiry No. 105

KPL's fluorescein, Texas-red or phycoerythrin labelled streptavidin offers high sensitivity and resolution in fluorescence microscopy. Texas-red and phycoerythrin-



Cy3 red, hydrophilic fluorescent label.



Dako CSA system.

labelled streptavidin can be used in single-colour analysis or in combination with other fluorophores for multiple labelling in flow cytometry and image analysis. AP- and HRP-labelled streptavidin are also offered. They are provided as liquid concentrates that are said to be stable for a minimum of one year.

Reader Enquiry No. 106

Equipment

Denley's Aquarius **microplate washer** incorporates an integral orbital shaker with the shaking speed automatically optimized according to the dispensed company volume. All washing cycles are user programmable through a simple four-key menu. A total of 99 wash programs can be named and stored, each having up to four linked protocols for more flexibility. Single or twin manifolds are available in 8- or 12-well formats. Moreover, the new bottle-management system allows multiple buffer usage with two 2-litre wash, the one 1-litre rinse and one 4.5-litre waste bottles, which can be switched electronically, with the 4.5-litre bottles utilizing liquid-level sensors. The software is designed to be easy to program and can handle complex programs containing up to four distinct washing protocols. For added security, the washer can be controlled from a PC and the wash procedure logged after each plate, if required.

Reader Enquiry No. 107

The new 20/20 research **luminometer** from Steptech is designed to offer improved sensitivity and versatility. It can measure light output in its specially designed sample compartment, which accommodates an 8 mm × 50-mm test tube, a 1.5-ml microfuge tube, a 28-mm scintillation vial and a 35-mm culture dish, as well as solids of various shapes and sizes. The instrument also accommodates wavelength-specific filters and a range extender. It is sensitive to a stated 0.1 femtograms of luciferase with a wide dynamic range of 10⁵ or greater. This can be extended to 10⁷ or greater with a range extending optical filter. In addition, an internal automatic injector is offered as an option for routine laboratory work, allowing the system to inject variable amounts of reagents (from 50–250 µl). The ASCII output can be transmitted through the instrument's RS232 port to any PC and it will interface with Epson-compatible printers. It also has a small footprint, measuring just 23.5 cm × 28 cm × 21 cm (L × W × H).

Reader Enquiry No. 108

BioRad has introduced a new small format **storage phosphor imager** with true chemiluminescence and radioisotope detection for quantification on a single phosphor imaging system. The GS-505 should provide an alternative to autoradiography as it is capable of detecting

all samples suitable for storage phosphor autoradiography, including chemiluminescent samples, without needing to alter the protocols. The system accepts standard small-format (20 × 25 cm) screens, bands and regions. When these lie close together they can be easily resolved on the GS-505 due to the instrument's high spatial resolution. The imager can qualify signals for samples over a five order-of-magnitude range. Available for use with



The GS-505 phosphor imaging system.

Power Macintosh, Macintosh and Windows 3.1/Windows95 platforms, the system is accessible to multiple users working on different platforms. The company offers a full range of compatible hardware, which can be combined with the system. This is designed to provide inexpensive documentation, visualization and quantification from a wide variety of fluorescent stains, including ethidium bromide, SYBR Green I/II, and SYPRO Orange. The GS-505 can also be upgraded easily to a GS-525, accepting both large and small screens.

Reader Enquiry No. 109

The InstantImager **electronic autoradiography system** from Packard Instruments is designed to be a faster way to retrieve quantitative images, achieving higher speeds through the use of Intel Pentium machines. The system comes complete with a customized computer, enabling the instrument to display and quantitate images in real time. The high-speed, digital-signal processing now takes place within a Pentium-based system, which includes 16 MB of RAM, InstantQuant for Windows and a gigabyte of hard drive space for archiving as many as 2,500 images.

Reader Enquiry No. 110

Literature

A new **diagnostic reagents directory** is now available from the Centre for Applied Microbiology and Research (CAMR). The directory provides guidance to the organization's reagents, services and expertise. Reagents in the directory include a range of antibodies and antigens developed as a result of CAMR's research into infectious diseases and vaccine development. There are three complementary aspects of the antibody range: stock antibodies, antibodies-to-order and bulk antibodies. Facilities are

also available for the production of bulk antibodies in quantities up to 20 g from user-supplied hybridomas. CAMR also has facilities with the highest levels of containment required to grow a wide range of bacterial and viral agents, should specialized antigens be required.

Reader Enquiry No. 111

The 1996–1997 Immunotech **immunoassay catalogue** features cytokine, retrovirology and histamine immunoassays. The catalogue introduces several new cytokine assays, including IL-7, IL-8, IL-10, GM-CSF and sGM-CSF receptor. Several other immunoassays are also highlighted, including cyclic AMP, cyclic GMP, neopterin, protein kinase and serotonin. The catalogue details each assay providing information on reagents, procedure, standard range, sensitivity, specificity and shelf life.

Reader Enquiry No. 112

These notes are compiled by Brendan Horton from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.

ADVERTISEMENTS

OLIGONUCLEOTIDES

£1 per base

- stringent quality control
- express turnaround
- column purified: ready to use

By research labs for research labs

Molecular Biology Services
King's College School of Medicine and Dentistry
Rayne Institute, 123 Coldharbour Lane, London SE5 9NU
Tel: 0171-346 3126 Fax: 0171-733 3877

READER ENQUIRY NO. 10

PACEMAKER®

Custom peptide, conjugate and antibody services

When your research depends on high quality at low cost turn to **PACEMAKER®**

All **PACEMAKER®** peptides, whether >70%, >80% or >95% purity, undergo hplc and mass spectrometric analysis.

Peptides are available in quantities from 1mg to 100mg and, with prices from **£115** per peptide and **£90** per conjugate, even the tightest budget is catered for.

Most importantly, you may come to us for advice. Based on many years experience and expertise we will try to find the best approach to fulfilling your custom peptide and antibody requirements.

Understanding the needs of the Life Scientist

AFFINITY

Research Products Limited

Mamhead Castle, Mamhead
Exeter EX6 8HD, United Kingdom
Tel +44(0) 1626 891010 Fax +44(0) 1626 891090

READER ENQUIRY NO. 70